Interference Characterisation, Location and Bandwidth Estimation in Emerging WiFi Networks

Arsham Farshad

Doctor of Philosophy

Institute of Computing Systems Architecture
School of Informatics
University of Edinburgh

2014

Abstract

Wireless LAN technology based on the IEEE 802.11 standard, commonly referred to as WiFi, has been hugely successful not only for the last hop access to the Internet in home, office and hotspot scenarios but also for realising wireless backhaul in mesh networks and for point-to-point long-distance wireless communication. This success can be mainly attributed to two reasons: low cost of 802.11 hardware from reaching economies of scale, and operation in the unlicensed bands of wireless spectrum.

The popularity of WiFi, in particular for indoor wireless access at homes and offices, has led to significant amount of research effort looking at the performance issues arising from various factors, including interference, CSMA/CA based MAC protocol used by 802.11 devices, the impact of link and physical layer overheads on application performance, and spatio-temporal channel variations. These factors affect the performance of applications and services that run over WiFi networks. In this thesis, we experimentally investigate the effects of some of the above mentioned factors in the context of emerging WiFi network scenarios such as multi-interface indoor mesh networks, 802.11n-based WiFi networks and WiFi networks with virtual access points (VAPs). More specifically, this thesis comprises of four experimental characterisation studies: (i) measure prevalence and severity of co-channel interference in urban WiFi deployments; (ii) characterise interference in multi-interface indoor mesh networks; (iii) study the effect of spatio-temporal channel variations, VAPs and multi-band operation on WiFi fingerprinting based location estimation; and (iv) study the effects of newly introduced features in 802.11n like frame aggregation (FA) on available bandwidth estimation.

With growing density of WiFi deployments especially in urban areas, co-channel interference becomes a major factor that adversely affects network performance. To characterise the nature of this phenomena at a city scale, we propose using a new measurement methodology called *mobile crowdsensing*. The idea is to leverage commodity smartphones and the natural mobility of people to characterise urban WiFi co-channel interference. Specifically, we report measurement results obtained for Edinburgh, a representative European city, on detecting the presence of deployed WiFi APs via the mobile crowdsensing approach. These show that few channels in 2.4GHz are heavily used and there is hardly any activity in the 5GHz band even though relatively it has a greater number of available channels. Spatial analysis of spectrum usage reveals that co-channel interference among nearby APs operating in the same channel can be a serious problem with around 10 APs contending with each other in many lo-

cations. We find that the characteristics of WiFi deployments at city-scale are similar to those of WiFi deployments in public spaces of different indoor environments. We validate our approach in comparison with wardriving, and also show that our findings generally match with previous studies based on other measurement approaches. As an application of the mobile crowdsensing based urban WiFi monitoring, we outline a cloud based WiFi router configuration service for better interference management with global awareness in urban areas.

For mesh networks, the use of multiple radio interfaces is widely seen as a practical way to achieve high end-to-end network performance and better utilisation of available spectrum. However this gives rise to another type of interference (referred to as *coexistence interference*) due to co-location of multiple radio interfaces. We show that such interference can be so severe that it prevents concurrent successful operation of collocated interfaces even when they use channels from widely different frequency bands. We propose the use of antenna polarisation to mitigate such interference and experimentally study its benefits in both multi-band and single-band configurations. In particular, we show that using differently polarised antennas on a multi-radio platform can be a helpful counteracting mechanism for alleviating receiver blocking and adjacent channel interference phenomena that underlie multi-radio coexistence interference. We also validate observations about adjacent channel interference from previous studies via direct and microscopic observation of MAC behaviour.

Location is an indispensable information for navigation and sensing applications. The rapidly growing adoption of smartphones has resulted in a plethora of mobile applications that rely on position information (*e.g.*, shopping apps that use user position information to recommend products to users and help them to find what they want in the store). WiFi fingerprinting is a popular and well studied approach for indoor location estimation that leverages the existing WiFi infrastructure and works based on the difference in strengths of the received AP signals at different locations. However, understanding the impact of WiFi network deployment aspects such as multi-band APs and VAPs has not received much attention in the literature. We first examine the impact of various aspects underlying a WiFi fingerprinting system. Specifically, we investigate different definitions for fingerprinting and location estimation algorithms across different indoor environments ranging from a multi-storey office building to shopping centres of different sizes. Our results show that the fingerprint definition is as important as the choice of location estimation algorithm and there is no single combination of these two that works across all environments or even all floors of a

given environment. We then consider the effect of WiFi frequency bands (e.g., 2.4GHz and 5GHz) and the presence of virtual access points (VAPs) on location accuracy with WiFi fingerprinting. Our results demonstrate that lower co-channel interference in the 5GHz band yields more accurate location estimation. We show that the inclusion of VAPs has a significant impact on the location accuracy of WiFi fingerprinting systems; we analyse the potential reasons to explain the findings.

End-to-end available bandwidth estimation (ABE) has a wide range of uses, from adaptive application content delivery, transport-level transmission rate adaptation and admission control to traffic engineering and peer node selection in peer-to-peer/overlay networks [1, 2]. Given its importance, it has been received much research attention in both wired data networks and legacy WiFi networks (based on 802.11a/b/g standards), resulting in different ABE techniques and tools proposed to optimise different criteria and suit different scenarios. However, effects of new MAC/PHY layer enhancements in new and next generation WiFi networks (based on 802.11n and 802.11ac standards) have not been studied yet. We experimentally find that among different new features like frame aggregation, channel bonding and MIMO modes (spacial division multiplexing), frame aggregation has the most harmful effect as it has direct effect on ABE by distorting the measurement probing traffic pattern commonly used to estimate available bandwidth. Frame aggregation is also specified in both 802.11n and 802.11ac standards as a mandatory feature to be supported. We study the effect of enabling frame aggregation, for the first time, on the performance of the ABE using an indoor 802.11n wireless testbed. The analysis of results obtained using three tools — representing two main Probe Rate Model (PRM) and Probe Gap Model (PGM) based approaches for ABE — led us to come up with the two key principles of jumbo probes and having longer measurement probe train sizes to counter the effects of aggregating frames on the performance of ABE tools. Then, we develop a new tool, WBest+ that is aware of the underlying frame aggregation by incorporating these principles. The experimental evaluation of WBest+ shows more accurate ABE in the presence of frame aggregation.

Overall, the contributions of this thesis fall in three categories — *experimental characterisation*, *measurement techniques* and *mitigation/solution approaches* for performance problems in emerging WiFi network scenarios. The influence of various factors mentioned above are all studied via experimental evaluation in a testbed or real-world setting. Specifically, co-existence interference characterisation and evaluation of available bandwidth techniques are done using indoor testbeds, whereas character-

isation of urban WiFi networks and WiFi fingerprinting based location estimation are carried out in real environments. New measurement approaches are also introduced to aid better experimental evaluation or proposed as new measurement tools. These include mobile crowdsensing based WiFi monitoring; MAC/PHY layer monitoring of co-existence interference; and WBest+ tool for available bandwidth estimation. Finally, new mitigation approaches are proposed to address challenges and problems identified throughout the characterisation studies. These include: a proposal for crowd-based interference management in large scale uncoordinated WiFi networks; exploiting antenna polarisation diversity to remedy the effects of co-existence interference in multi-interface platforms; taking advantage of VAPs and multi-band operation for better location estimation; and introducing the jumbo frame concept and longer probe train sizes to improve performance of ABE tools in next generation WiFi networks.

Acknowledgements

I would like to express my special appreciation and thanks to my supervisors Mahesh Marina and Francisco Garcia for their supervision and support throughout my PhD. Mahesh treated me as his son by giving priceless advice and support. I was also be lucky to have Francisco as my second supervisor; he has always encouraged me to continue with his feedback and humble attitude. I would also like to thank my previous supervisor, Siavash Khorsandi, who has encouraged and help me to pursue PhD.

I would also want to thank my exam board: Paul Patras, Robert Atkinson and Bjoern Franke for their comments and suggestions; special thanks to Paul for his detailed comments that considerably improved the quality of the thesis. I would also like to specifically thank Harald Haas and Myungjin Lee for their advice and help in different parts of the research reported in this thesis.

I would also like to thank my colleagues and friends: Giacomo Bernardi, Sofia Pediaditaki, Lito Kriara, Valentin Radu, Matt Calder, Manik Rathinakumar and Jiwei Li.

A special thanks to my family; their prayer for me was what sustained me so far. I should also thank my English teacher Mrs. Mackay for her kindness and support.

Last but not least, I would like to express appreciation to my beloved wife Zoha who has endured spent sleepless nights and was always my support at times when there was no one to answer my queries.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified. Some of the material used in this thesis has been published in the following papers:

- A. Farshad, M. Lee, M. K. Marina and F. Garcia, On the Impact of 802.11n Frame Aggregation on Available Bandwidth Estimation, has been accepted as a full paper for presentation at the IEEE International Conference on Sensing, Communication, and Networking (SECON), Singapore, June-July 2014.

- A. Farshad, M. K. Marina and F. Garcia, Urban WiFi Characterization via Mobile Crowdsensing, has been accepted as a full paper for presentation at the 14th IEEE/IFIP Network Operations and Management Symposium (NOMS), Krakow, Poland, May 2014.

- A.Farshad, J. Li, M. K. Marina and F. Garcia, A Microscopic Look at WiFi Fingerprinting for Indoor Mobile Phone Localization in Diverse Environments, in the Proceedings of 4th International Conference on Indoor Positioning and Indoor Navigation (IPIN), Montbeliard - Belfort, France, October 2013.

- A. Farshad, M. K. Marina and F. Garcia, Experimental Investigation of Coexistence Interference on Multi-Radio Platforms, in the Proceedings of 10th Modeling and Optimization in Mobile, AdHoc and Wireless Networks (WiOpt), 2012.

(*Arsham Farshad*)

To my wife, Zoha

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Emerging new applications such as high definition video on demand, IP-TV and cloud-based applications together with new technologies based on the Internet such as Internet of Things (IoT), smart grids, smart home, smart cities and smart transportation introduce explosive amount of data to transfer to/from the Internet [6] (it is estimated that we will have 16 billion connected devices by 2020 which shows at least two devices per person on earth, on average [7]). As WiFi communication is the most popular last-hop technology to connect to the Internet, it is the first and cheapest candidate to accommodate communication of such high volume of data.

In comparison with mobile cellular technologies like GSM and UMTS, WiFi has significant advantages from operating in unlicensed spectrum and cheap hardware, which in turn motivates offloading cellular traffic to available WiFi networks.

However, performance of applications and services running over WiFi networks can be affected (un)favourably by several underlying factors including interference, multi-band operation, use of virtual access points (VAPs), CSMA/CA channel access protocol, spatio-temporal channel variations, protocol overhead and MAC/PHY layer enhancements.

Below we discuss some of these factors relevant to this thesis; a more detailed description is provided in Chapter 2.

Co-Channel Interference:

Among the aforementioned factors, interference is a key source of performance degradation in both legacy WiFi networks and emerging WiFi networks. Among different types of interference, *co-channel interference (CCI)* or mutual interference is

the most common one.

Co-channel interference concerns the use of a common channel by several nearby and mutually interfering transmitters, and results from un-coordinated and un-planned WiFi networks operating in close proximity of each other and using the same channel. For example, if two wireless broadband routers at two neighbouring flats use the same channel (say, channel 6) then that causes CCI, effectively reducing the throughput achievable by wireless devices associated with either of those routers.

Authors in [8] experimentally investigate the effect of AP density on performance of different applications such as web and multimedia using the ORBIT testbed [9]. Results of this study show that in an unplanned multi-AP deployment scenario, increasing the number of APs causes a significant increase in collision rate and consequently a high drop in throughput. They reported 50% drop in throughput with only four interfering APs. Additionally, media streaming performance is reported to take a big hit in the presence of co-channel interference. For the voice over IP (VoIP) application, substantial performance degradation is seen, with just three APs — average latency increases from 54ms in the single AP scenario to 304ms in the scenario with four uncoordinated APs; jitter also increases four-fold with the multi-AP scenario.

Interference from non-802.11 networks and devices (*e.g.*, Bluetooth, ZigBee, microwave ovens, baby monitors, cordless phones) is another source of co-channel interference that affects wireless transmissions in 2.4GHz band. Authors of [10] experimentally study the mutual effect of concurrent transmissions of low power interferers like Bluetooth and ZigBee transmitters and WiFi. Their results show that ZigBee and Bluetooth devices operating near WiFi devices suffer considerably with degradation in traffic performance by 41% and 68%, respecitvely; on the other hand, WiFi performance is hardly affected by the other two. However, high power non-802.11 devices such as surveillance cameras, cordless phones and baby monitors have a devastating effect on the performance of a WiFi link. They can cause complete loss of connectivity as reported in [11].

Coexistence Interference:

In addition to the co-channel interference from the co-located networks, multi-radio networks also suffer from coexistence interference. In multi-radio networks, individual APs/routers are equipped with multiple radio interfaces, each configured to a different channel. Advantages of such an architecture for achieving high end-to-end network performance, keeping (co-channel) interference low and better utilization of available spectrum are now well established.

Multi-interface routers were initially proposed to improve the performance of the mesh networks. Since mesh networks have evolved as multi-hop extensions of wireless LANs (WLANs), they typically use radio interfaces based on IEEE 802.11. More recently, as AP densities grew and user demand for greater bandwidth increased, multi-radio APs emerged for improved interference management and spectrum utilization. For example, Cisco WAP561 Wireless-N Dual Radio [12] APs are equipped with two radio interfaces that can work simultaneously if they are setup in channels from different bands (*i.e.*, one in 2.4GHz and one in 5GHz).

Realizing the benefits of multi-radio architectures in practice poses challenges that are often abstracted out in simulation based evaluations. These concern interference resulting from collocation and simultaneous operation of multiple radio interfaces within a multi-radio platform. Intuitively, adding a second wireless interface configured to use a different orthogonal channel should double the capacity. However, observation in different previous researches (*e.g.*, [13, 14]) show that the newly added interface does not actually double throughput; instead a transmitting interface appears as a high power noise to co-located receiving interface, leading to lower throughput than what would be achieved with a single interface. In fact, the authors of [14] report 60% UDP throughput degradation due to the coexistence of multiple interfaces.

Following [15], we will refer to this interference as *multi-radio coexistence interference*. This interference is described by Zhu et al. [15] as a composition of three phenomena: *receiver blocking*, *intermodulation* and *transmitter noise*. While receiver blocking and inter-modulation are results of the limited dynamic range of output power amplifier and A/D converter, transmitter noise or adjacent channel interference (ACI) refers to the out-of-band emission seen by receivers in close proximity of a transmitter — due to imperfect filtering at the transmitter antennas.

To mitigate effects of ACI, increasing distance between channels in frequency domain and physical separation of antennas are proposed; for example, [14] proposes to have 80MHz (3 non-overlapping channels) separation between centre frequency of two chosen channels and minimum of 40cm separation between two antennas. These proposals are unsatisfactory because they lead to bulky platforms (due to large antenna separation) or cause inefficient spectrum use.

Spatio-Temporal Channel Variations:

Interference is induced by transmission of other 802.11 or non-802.11 devices, while its effect may be more or less severe depending on spatio-temporal channel variations. Spatio-temporal channel variations happen due to the changes in the signal

propagation environment characteristics. WiFi signal propagation is mainly influenced by the environment's material, floor plan and furniture of a building. Signal coverage map changes over time due to people movement and their effects on the signal propagation environment.

Spatial channel variations is exploited by WiFi fingerprinting based localisation approaches. Precisely, it helps to have distinctive signal fingerprints for different geographical locations. In new generation of 802.11 devices based on the 802.11n/ac, spatial diversity and spatial multiplexing are used to increase robustness and capacity, respectively. While the spatial channel variation can be exploited to achieve benefits such as those mentioned above, temporal variations tend to be mostly harmful. These induce variability in the wireless link's capacity and available bandwidth, even when there is no contending or interfering traffic; temporal channel variation appears as measurement noise in WiFi fingerprints.

Virtual Access Points (VAPs):

In multi-interface platforms, for each logical WiFi interface seen by the system there is a corresponding physical WiFi interface; in contrast, VAPs enable to define different logical WiFi interfaces by time sharing resources of a physical interface. Having multiple logical WiFi interfaces allows establishing different WLANs. For example, an AP equipped with a single physical WiFi interface, can advertise two different WLANs with different network names(*e.g.*, NetworkA and NetworkB) and different supported features (*e.g.*, open access and WPA2).

VAPs are introduced to realise the different security policies for different user groups; VAPs are to WiFi networks just as Virtual LANs (VLANs) are to switched Ethernet networks. Initially, the VAP functionality was introduced in high-end business class APs; later, it was also introduced in lower class home and small office class of products inspired by high demand for sharing home Internet access for guest users or as a hotspot zone (*e.g.*, BTFON WiFi services [16]). APs supporting VAPs are usually pre-configured to two virtual WiFi networks: one is for the guest users with only access to the Internet. Another VAP advertises itself with a different name and pass phrase and provides a full access to all network resources (*e.g.*, networked printer, shared storage) for the associated users.

Through our urban WiFi deployment characterisation, we find that VAPs' deployment are very common in public/office buildings such as university buildings, shopping centres and hospitals. For example, in one shopping centre, 8 VAPs are advertised belonging to one physical AP by transmitting eight different network names in

different management beacon frames, all within a 100ms period. Though the main purpose of introducing VAPs for providing different security access levels is generally well-known, their effects on performance of applications and services like WiFi based location estimation services have not been thoroughly studied yet.

Multi-Band Operation:

Some emerging WiFi standards like IEEE802.11ac only supports transmission in the 5GHz band, suggesting that 5GHz band will become more popular for WiFi communications in the future especially for high throughput applications.

More channels in the 5GHz band, less amount of co-channel interference, and lower level of noise from non-802.11 devices are among other motivations to switch to the 5GHz band. Specifically, 19 orthogonal (non-overlapped) channels from the 5GHz band are available to use in UK [17], whereas only three non-overlapped channels are available in the 2.4GHz band. A larger number of available channels also decreases the probability of co-channel interference and offers more freedom in frequency management. Moreover, transmission in the 5GHz band is more prone to non-802.11 devices such as cordless phones, monitoring cameras and microwave oven, usually used in close vicinity of WiFi devices. Although, transmission on some channels of the 5GHz band should be aware of the primary users of this band, *e.g.* military and weather radars, by implementing dynamic frequency selection (DFS) [3].

However, where communication range is important, transmitting on 2.4GHz band is preferred as it results in longer range compared to transmitting in the 5GHz band. Introduction of WiFi interfaces which support multi-band operation, APs equipped with dual interfaces configured to operate on both bands simultaneously and existence of legacy APs/devices which only support transmission in 2.4GHz band, all imply that future WiFi networks will have transmissions on both bands with greater use of channels of the 5GHz band.

Differences in the amount of interference along with differences in spatio-temporal behaviour between 2.4GHz and 5GHz bands not only affect communication performance (*i.e.*, throughput and packet loss) but also might implicitly affect the performance of services like WiFi fingerprinting, which has not been thoroughly explored yet.

MAC/PHY Layer Enhancements:

Underlying factors such as interference and spatio-temporal channel variations are induced by either transmission of other devices or changes in the propagation environment due to user mobility. However, MAC/PHY layer enhancements are of a different

kind that are introduced to improve performance and reliability. New enhancements in both physical and MAC layer are introduced in 802.11n and its successor 802.11ac standards that considerably improves communication performance and connection reliability in comparison to the legacy 802.11a/b/g standards.

A key enhancement in the PHY layer is MIMO or using multiple input and multiple output. The majority of 802.11n hardware is equipped with two physical transmit circuits connected to two antennas supporting 2×2 communication. Two main introduced features based on MIMO are Spatial Division Multiplexing (SDM) and Space-Time Block Coding (STBC). SDM exploits the multi-path feature of signal propagation to transmit multiple different streams carrying different bits of data to a destination or even multiple destinations in 802.11ac. STBC, however, enhances reliability by transmitting coded blocks of same data in multiple streams to increase the chances of retrieving corrupted data at the receiver. Other enhancements in the physical layer are to use a wider channel by bonding channels of 20MHz (maximum two channels for 802.11n and upto eight channels in the 802.11ac standard), shorter guard interval between OFDM symbols (SGI) and higher transmission bit-rates with a denser modulation and coding schemes.

Translating high data rates from PHY layer enhancements into corresponding high application performance requires keeping the inefficiencies due to protocol overheads low. Accordingly, 802.11n/ac includes two key enhancements for reducing MAC layer protocol overhead — Frame aggregation (FA) and block acknowledgements— FA improves the efficiency by transmitting back-to-back frames wrapped in a big frame upon each successful channel access. Successful receipt of individual frames within each aggregated frame is reported together via one block ACK frame. The impact of these two MAC layer enhancements in comparison with other PHY features has been experimentally evaluated in [18]. The paper shows that without enabling frame aggregation, the maximum UDP throughput cannot exceed 110Mbps with a 2×2 MIMO WiFi interface, whereas the achievable throughput goes up to 240Mbps with frame aggregation.

While the intention behind the above mentioned enhancements is to improve higher layer performance, that may not always be the outcome especially when higher layer applications and services are unaware of their use. We present one such instance in this thesis in our study of available bandwidth estimation (ABE) over 802.11n networks.

1.2 Thesis Contributions

1.2.1 Common Theme: Experimental Characterisation

Experimental characterisation along with simulation and analytical modelling are major approaches for evaluating the performance of a protocol, system or network. Each approach, though, has its own advantages and disadvantages.

While the simulation and analytical modelling approaches have several advantages (flexibility in configuration, fast evaluation, repeatable results, scalability, etc.), in the context of wireless networks several studies (*e.g.*, [19, 20, 21]) have reported significant gaps between results from testbed experiments or real-world evaluation and those obtained via simulation. This is mainly due to the unrealistic nature of radio signal propagation models used and abstracting some key implementation aspects of protocols.

In this thesis, we undertake experimental characterisation of the impact of factors discussed in the previous section in the following different contexts:

- *Urban WiFi characterisation via mobile crowd sensing (Chapter 3):* characterising co-channel interference in a city scale.
- *Coexistence interference characterisation and mitigation in multi-radio platforms (Chapter 4):* characterising coexistence interference in multi-radio 802.11-based platforms.
- *Indoor WiFi fingerprinting based localisation in diverse environments (Chapter 5):* characterising the influence of spatio-temporal channel variations, VAPs and multi-band operation on the WiFi fingerprinting based location estimation service.
- *Available bandwidth estimation over new generations of WiFi networks (Chapter 6):* characterising effects of MAC/PHY layer enhancements, especially frame aggregation (FA), on available bandwidth estimation (ABE) over 802.11n networks.

This thesis also contributes new measurement techniques and mitigation/solution approaches. It introduces mobile crowdsensing based monitoring of urban scale WiFi deployments and suggests a way to use the monitoring information for cloud based adaptive channel selection for individual WiFi networks. It also proposes a monitoring

approach for direct observation of MAC/PHY layer behaviour in presence of the adjacent channel interference (ACI) by leveraging MAC/PHY layer registers designed for Clear Channel Assessment (CCA). This measurement approach is used to show benefits of exploiting antenna polarization diversity, the approach proposed in this thesis to alleviate effects of coexistence interference. Moreover, we propose a new measurement tool, WBest+, for estimating available bandwidth over next generation WiFi networks based on the 802.11n standard.

1.2.2 Urban WiFi Networks Characterisation Via Mobile Crowdsensing

Co-channel interference is identified as one of the main factors that hurts performance of WiFi communication. It is common to see several neighbouring networks sharing a channel in un-planned WiFi networks such as home and small office WiFi networks. Understanding the extent of such channel sharing (in other words, the severity of co-channel interference) is useful to devise adaptive mechanisms to mitigate such interference but this requires a monitoring infrastructure that can continuously observe deployments of WiFi networks and a spectrum management engine that analyses information from monitoring system and identifies suitable channels to use for each observed network. *Setting up a fixed monitoring infrastructure* and *wardriving* are two proposed approaches in the literature for urban-scale WiFi monitoring. Like Argos [22], fixed monitoring networks are implemented by deploying several sensors (sniffers) which are deployed on rooftop of buildings or street light across a city. Success of infrastructure based monitoring network relies on different factors including capabilities of sensors, scalability of the network and operation cost. For example, in Argos, the deployed sensors (sniffers) have abilities to capture and analyse WiFi traffic. Analysed traffic information is exploited to detect malicious traffic, classify wireless access points and clients, and track the mobility of WiFi-equipped public transport vehicles. However, setting up and operating such networks is very expensive and tedious. In large cities, which are rapidly expanding and changing, scalability of the monitoring network is a real and serious challenge.

An alternative approach to setting up a monitoring network is wardriving [23] using a sensor (it can be as simple as a laptop with wireless interfaces) installed in a car or carried by a person to explore the city and capture information of wireless network deployments while moving around. The information captured in one run of wardriving

may become outdated very soon and repeating city scale measurement is an expensive task that makes wardriving not suitable for adaptive spectrum management purposes.

To answer shortcomings of traditional approaches, we propose a *mobile crowdsensing* approach using the WiFi interface of smartphones as a measurement sensor and leveraging participants with smartphones to carry out measurement in their surrounding environments. When a group/community of participants (a *crowd*) is engaged with suitable incentives, mobile phone sensing becomes even more compelling for continual and fine-grained spatio-temporal monitoring of the phenomenon of interest in a *cost-effective* manner. One of the advantages of the mobile crowdsensing approach compared to wardriving and fixed monitoring infrastructure is its flexibility for naturally extending the scope of measurements to inside of buildings where the majority of APs would be actually deployed. However, limited battery capacity, storage of the sensing data and cost of mobile data plan are common barriers in mobile crowd sensing. Using wise measurement techniques such as that introduced in [24] can lessen the side effects of crowdsensing approaches.

To validate the applicability of our proposed mobile crowdsensing approach and as a proof-of-concept, we first compare the effects of deploying smartphones against laptops usually used in the wardriving. The results from the experimental validation demonstrate that mobile crowdsensing with commodity smartphones can yield similar results to those obtained via carefully conducted wardriving campaigns.

Backed by these results, we conduct a measurement study in the city of Edinburgh using Android phones which feature a WiFi interface supporting both 2.4GHz and 5GHz unlicensed bands (*e.g.*, Samsung Galaxy S3 ([25]). We rely solely on passive scanning based measurements, listening to AP beacons and leveraging participants with mobile phones travelling on public transport buses. Then, we analyse the information collected by smartphones to study and map the spatial distribution of APs; their operating frequency channels and locations that might suffer from co-channel interference.

We also expand the scope of monitoring to inside of buildings and compare our observations from outdoor to what we observed from indoor for different buildings in different parts of the city. This study illustrates similarity between indoor and outdoor observations.

Last but not least, our measurement study reveals that co-channel interference is very common in urban areas owing to existence of many WiFi networks which are deployed in close vicinity without considering how neighbouring APs are sharing the

available spectrum. This can be seen as a real world evidence to show that vast research on self-organisation mechanisms for channel and transmit power allocation in unplanned WiFi deployments (e.g., [26]) has not actually materialized. We observe that the impediment for large-scale deployment of intelligent self-organisation mechanisms in practice may not be technical but rather the lack of market incentives for their application. With this in mind, we outline an alternative approach that may find greater real world acceptance. The idea is for a mobile crowdsensing based urban WiFi monitoring system to continually feed spectrum usage measurements to a cloud based back-end, which takes the global awareness of spectrum usage and interference conditions to determine the best channel for each participating WiFi AP (*e.g.*, home WiFi router).

Our findings and contributions are summarised as follows:

- We propose a mobile crowdsensing based measurement approach for monitoring large scale un-planned WiFi networks. We obtain similar qualitative results in comparison to a carefully done wardriving study and observe similar results in agreement with other previous studies following different measurement approaches, thereby validating our proposed approach.

- We find that observations about WiFi deployments in public areas of several different indoor environments match that of WiFi deployment characteristics at city-scale.

- WiFi spectrum usage is quite unevenly distributed across 2.4GHz and 5GHz unlicensed bands as well as among various channels within the 2.4GHz. Many WiFi APs contend on the same channel with around 10 other APs (and their clients) in the nearby vicinity, thereby potentially experience severe interference. This is a result of the common practice of uncoordinated and non-adaptive channel assignment to home WiFi routers which are often left to use preset factory configuration settings for channels.

- We also look into the distribution of open APs, which could be leveraged for vehicular WiFi access [27]. We observe that the availability of open APs along contiguous road segments is limited to few parts near the city centre.

- We outline a cloud based WiFi spectrum management service for WiFi APs in urban areas (e.g., home wireless routers) that can make use of results from mobile crowdsensing based urban WiFi monitoring for better interference management.

This work has been accepted for presentation as a full paper at the *14th IEEE/IFIP Network Operations and Management Symposium (NOMS), Krakow, Poland, May 2014.*

1.2.3 Coexistence Interference Characterisation and Mitigation

In addition to the co-channel interference from the co-located networks, multi-radio networks suffer from coexistence interference. Multi-radio architectures are introduced to help achieve high end-to-end network performance and reduce co-channel interference by more efficient spectrum usage. Realizing these benefits in practice, however, poses challenges in the form of coexistence interference, resulting from co-location and simultaneous operation of multiple radio interfaces within a multi-radio platform.

We study the effects of coexistence interference in the context of multi-radio mesh networks and show that performance degradation due to coexistence interference can be significant if adequate care is not taken to mitigate it. Our study also represents multi-interface APs recently proposed for deployment in infrastructure WLANs.

Specifically, we study the impact of a transmitting interface on the reception over a collocated interface and consider receiver blocking and transmitter noise/adjacent channel interference (ACI) effects. In contrast to Zhu et al. [15] who focused on the case of platforms with multiple radios using heterogeneous wireless technologies (e.g., Bluetooth, WiFi, WiMax), our focus is on router/AP platforms with multiple 802.11 radios in concurrent operation. Physical separation of antennas on 802.11 multi-radio platforms and separation between channels used by different collocated radio interfaces are common ways to avoid performance degradation due to multi-radio coexistence interference [13, 28, 29, 14, 30, 31].

The required amount of antenna and channel separation varies depending on factors such as transmission power and bit-rate. Ideally, we would want these separations to be as small as possible while allowing to transmit at maximum bit-rates supported by link/channel quality and at transmit power up to the regulatory limit. Small antenna separation will lead to a compact platform that is easier and cheaper to deploy, an especially important consideration for indoor WLAN/mesh deployments. Smaller channel separation can potentially allow better utilization of available spectrum.

To relieve the need for increased antenna/channel separation without limiting transmit power/bit-rate, we consider *antenna polarization*, for the first time, as an extra knob

to introduce an additional coupling loss, drop in strength of interference, up to 20dB between collocated antennas.

Experimental results show that having differently polarised antennas can reduce antenna separation five times lower than with identically polarised antennas (6cm vs. 30cm). The required channel separation to prevent ACI is reduced and possible transmission bit-rates increase by upto four times compared to the identically polarised antenna setup. Mentioned benefits do not come at the cost of network connectivity. Using measurements from an actual indoor multi-radio mesh network testbed, we show that the use of differently polarised antennas has a small effect on the mesh network topology and link qualities.

We also introduce a new ACI measurement technique that involves reading the channel state registers on the MAC/PHY layer; these are designed for Clear Channel Assessment (CCA)— a procedure carried out by each station before transmitting to ensure that the channel is not occupied by others [3]. Using this technique shows that when one interface is transmitting, the MAC on the other collocated interface using an adjacent but non-overlapping channel senses the channel to be BUSY. By increasing the channel separation between two interfaces, the percentage of time that MACs sense their operating channels as BUSY reduces.

Our findings and contributions are summarised as follows:

- Experimentally show the severity of multi-radio coexistence interference on the traffic performance due to receiver blocking even when using different frequency bands.

- Demonstrate the use of antenna polarization as a means to alleviate multi-radio coexistence interference with little negative side effects.

- Characterise adjacent channel interference using direct observation of MAC behaviour and validation of observations from prior work based on indirect packet delivery ratio measurements (*e.g.*, [14]).

This work is published in *the Proceeding of 10th Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2012.*

1.2.4 Indoor WiFi Fingerprinting Based Localisation in Diverse Environments

For location-aware applications and IoT, knowing the current location of an object/thing or people becomes much more important. For outdoor localisation, GPS is now a commodity whereas it is shown to be unreliable and energy hungry when used indoors [32]. On the other hand, WiFi fingerprinting has emerged as a popular WiFi based localisation technique for indoor use in the past 10-15 years since the idea was first put forth in the RADAR system [33]. The attractive thing about the WiFi based localisation approach is that it exploits the prevalent WiFi infrastructure in many indoor environments and the presence of the WiFi interface is now very common in handheld devices like smartphones and expected to be so in several other objects/things in future. The main idea behind WiFi fingerprinting is to use signal characteristics at each location (usually signal strength from visible APs) as a signature to infer location. We take a microscopic look by studying various aspects underlying the WiFi fingerprinting approach in the context of indoor mobile phone localisation. Different aspects considered in our study include: the definition of a fingerprint, run-time location estimation algorithms, frequency band and presence of virtual access points (VAPs).

Our investigation considers several different real indoor environments ranging from a multi-storey office building to shopping centres of different sizes to study the spatial effects of signal propagation. Eight different definitions of fingerprints are considered to capture both spatial and temporal effects of signal propagation. We consider fingerprint definitions that are defined based on the RSSI, AP visibility and combinations of both. With respect to location estimation algorithms, we compare three different deterministic techniques (including the often used Euclidean distance based nearest neighbour method) with two probabilistic techniques that use Gaussian and Log-normal distributions for RSSI modelling.

We obtain WiFi fingerprinting data for our study using a custom mobile application specifically developed for this purpose and installed on an Android phone. Results show that deterministic location estimations (*e.g.*, Manhatan, Mahalanobis) outperform probabilistic schemes across different environments. However, the choice of fingerprint definition has as much or more impact than the location estimation algorithm. The best choice of fingerprint definition and location estimation algorithm are also different between different environments, even between floors of a building.

From our study on urban WiFi characterisation, we found that operation in both

2.4GHz and 5GHz bands is common in indoor environments. The difference in signal propagation characteristics between these two bands motivates us to study potential effects of multi-band operation on the location estimation and WiFi fingerprint definition.

To study the impact of multi-band operation, we use a mobile phone that supports both 2.4GHz and 5GHz bands (*e.g.*, Samsung Galaxy S3 [25]) to carry out localisation based on three training data sets: 2.4GHz only, 5GHz only and both 2.4GHz and 5GHz. Results show that there is a significant benefit of having APs working in the 5GHz band. Our analysis to identify the reasons behind this result shows that it is due to lower level of co-channel interference in the 5GHz band given the larger number of orthogonal channels available. This results in lower beacon loss and thus in a lower temporal variation of the received signal strength. We also consider the effect of virtual access points (VAPs), which is another aspect found to be common from our urban WiFi characterisation. Contrary to intuition, we find that the presence of VAPs significantly improves WiFi fingerprinting accuracy which we believe is due to two reasons: VAPs have a substantial influence on the AP density, a factor known to affect accuracy with WiFi fingerprinting; and fingerprints obtained from different co-located VAPs operating on the same channel are somewhat dissimilar, capturing the temporal variability inherent to wireless signal propagation and providing robustness against it.

Our findings are summarised as follows:

- Our analysis shows that the fingerprint definition is at least as important as the choice of location estimation algorithm and there is no single combination of fingerprint definition and localisation algorithm, that always yields the optimum localisation result across all the different environments. Even different floors within the same building have different optimum combination of fingerprint definition and localisation algorithm.

- We consider the impact of frequency band used (2.4GHz vs. 5GHz) on WiFi fingerprinting and find that 5GHz offers relatively better location accuracy due to lower level of beacon loss because of the co-channel interference.

- We find that the presence of VAPs significantly improves WiFi fingerprinting accuracy.

This work is published in *the Proceedings of 4th International Conference on Indoor Positioning and Indoor Navigation (IPIN), Montbeliard - Belfort, France, Oct*

2013.

1.2.5 Available Bandwidth Estimation over Next Generation WiFi Networks

End-to-end available bandwidth estimation (ABE) approaches, typically based on active measurement with probing packets, have received considerable attention in the literature during recent years given their importance for applications such as: adaptive content delivery (*e.g.*, media servers and adaptive media streaming [34, 35, 36, 37]), admission control, traffic engineering, peer node selection in peer-to-peer/overlay networks [1, 2]. As accessing the Internet from wireless enabled devices (*e.g.*, laptops, tablets and smartphones) becomes very popular, the need for available bandwidth estimation over paths with wireless links — especially WiFi links — takes greater importance given that people spend most of their time indoors where WiFi is prevalent (*e.g.*, homes). While there exists a substantial body of work (*e.g.*, [38, 39, 40, 41, 42, 43, 44]) examining available bandwidth estimation over path with 802.11 wireless links, these only consider the case of legacy 802.11 networks (802.11a/b/g).

The enhancements of PHY/MAC layers introduced in the new generation of 802.11 networks based on the 802.11n/ac standards increase the achievable bandwidth to more than 1Gbps. While enhancements such as the use of MIMO and wider channels are responsible for increased physical layer bit-rates in 802.11n and 802.11ac, MAC layer enhancements especially *frame aggregation (FA)* are key to translating those bit-rates to higher throughputs above the MAC layer [45]. Not all proposed enhancements are mandatory to support; for example in 802.11n supporting channel bonding, short guard interval, STBC and $3 \times 3/4 \times 4$ MIMO are optional features.

In our study of ABE over 802.11n networks, we first find that FA is the most dominant feature of 802.11n standard that adversely affects ABE. On the other hand, supporting aggregation is mandatory in 802.11n/ac standards and FA was also shown to have a big impact in maximising application layer throughput [18]. Given these two observations, we focus our study on characterising the effect of FA on the performance of commonly used ABE approaches.

Taking into account the fact that ABE tools mostly take one of two approaches, the Probe Gap Model (PGM) or the Probe Rate Model (PRM) [2, 44], we experiment with three ABE tools — WBest [40], DietTopp [39] and pathChirp [46] — that represent PGM and PRM approaches, and also have been previously evaluated in the legacy

802.11a/b/g context. We study their available bandwidth estimation accuracy in the presence of 802.11n frame aggregation across a wide range of cross-traffic scenarios.

Results illustrate that all tools underestimate available bandwidth in presence of FA with PRM-based tools like DietTopp and pathChirp having lower errors. The amount of estimation error is dependent on the amount of cross traffic transmitted with measurement traffic. For example, by enabling FA and with no cross traffic, estimation error is 20% for DietTopp and 53% for WBest. However, increasing the cross traffic increases estimation errors to 60% for DietTopp and 100% for WBest.

To understand why frame aggregation has such a detrimental effect on the accuracy of ABE tools and as a tool for experimenting mechanisms aimed at improving ABE performance, WBest is chosen for further study given it is a faster tool compare to PRM-based tools, a crucial feature to apply in multimedia applications (*e.g.*,media servers).

Our in-depth analysis shows that FA can result in very low or very high probe dispersion times depending on whether or not measurement probes get aggregated, and directly using those dispersion time values will lead to under/over-estimations. We also find that to capture the effect of FA on available bandwidth we need to use larger number of probes. Given these findings about the behaviour of WBest in the presence of FA, we devise a new tool, WBest+, an improved variant of WBest. WBest+ treats aggregated probes as a jumbo probe for calculating the dispersion time and transmits a larger number of probes to better capture FA effects. Experimental results show that WBest+ outperforms WBest in both link capacity and available bandwidth estimation. Our findings are summarised as follows:

- We consider available bandwidth estimation in the 802.11n context for the first time, and show that frame aggregation not only is a key feature from a performance viewpoint but also is the most dominant feature among 802.11n features affecting ABE.

- Our comparison of different ABE tools in various cross-traffic scenarios focusing on the impact of frame aggregation leads to the following observations:

    1. The frame aggregation feature significantly hurts the accuracy of all considered ABE tools.

    2. DietTopp and pathChirp, the tools that follow the Probe Rate Model (PRM), are relatively more robust in presence of frame aggregation compared to WBest that belongs to the Probe Gap Model (PGM) category.

3. The tools under test tend to under-estimate the available bandwidth in presence of FA across many different cross-traffic rates in all contending/interfering scenarios considered in this study.

- We introduce a new tool, WBest+, based on the WBest tool. It incorporates our new introduced principles of jumbo probes and a larger number of probes. WBest+ shows better performance in achieving robust and accurate ABE over FA-enabled 802.11n networks. Our approach can be generalized for new 802.11ac standards, as FA is specified as a core and mandatory feature to support.

This work has been accepted for presentation as a full paper at the *IEEE International Conference on Sensing, Communication, and Networking (SECON), Singapore, June-July 2014.*

1.3 Structure

The rest of this thesis is organised as follows.

Chapter 2 gives an overview of the technologies behind emerging networks discussed in the thesis. It reviews concepts, terminology and algorithms used in the subsequent chapters and discusses related work published in the literature for each subject under study.

Chapter 3 introduces our crowdsensing measurement approach for characterising WiFi deployments in urban areas and presents measurement results using the proposed approach to characterise the WiFi deployment for city of Edinburgh. Results include the WiFi spectrum usage map and the potential interference spots map. The idea of cloud based approach based on the crowdsensed monitoring is also discussed for adaptive spectrum management of uncoordinated WiFi networks.

In Chapter 4, we study another type of interference, coexistence interference, that hurts the effectiveness of multi-radio platforms. We employ a new approach to measure ACI effects via direct observation of MAC/PHY layer behaviour. To lessen the effect of receiver blocking and ACI, antenna polarization diversity is exploited; we show for indoor mesh networks that differently polarized antennas have any noticeable negative effect on network connectivity while it enables more efficient channel assignment and much more compact nodes, which promotes multi-radio platforms suitable for indoor deployment and other applications (*e.g.*, IoT).

The indoor WiFi characterisation study in Chapter 3 also revealed that the deployment of VAPs and multi-band operation is fairly common in office buildings and shopping centres. This, along with the importance of knowing the physical location and prevalence of location-aware applications, motivate us to study the effects of such features on the performance of location estimation algorithms and the characteristics of WiFi fingerprinting definitions. Specifically, in Chapter 5, different WiFi fingerprinting definitions and location estimation algorithms are evaluated across different buildings and their results are presented.

Motivated by the popularity of bandwidth hungry applications demanding very high throughput (*e.g.*, high definition video streaming), new enhancements are introduced in 802.11n/ac standards. In Chapter 6, we study the effect of frame aggregation, the main MAC layer enhancement introduced in 802.11n standard, on the performance of ABE tools. A new tool, WBest+, is introduced in this chapter, which improves the accuracy of ABE in presence of FA.

In Chapter 7, we present overall conclusions of this thesis and discuss future work.

Chapter 2

Background and Related Work

In §2.1, background information about emerging WiFi networks are given. Specifically, multi-radio mesh networks, virtual WiFi networks and features of high throughput networks based on the 802.11n standard are introduced. Interference as one of the main factors affecting communication performance is discussed in §2.2; co-channel interference (§2.2.1) and coexistence interference (§2.2.2) are introduced. Previous studies in coexistence interference characterisation and mitigation are reviewed. Antenna polarisation is explained in §2.3; then later in Chapter 4, antenna polarisation diversity is exploited to mitigate coexistence interference. City scale monitoring approaches include: deployment of a fixed-infrastructure monitoring network, wardriving, and crowdsensing, which are studied in §2.4. WiFi fingerprinting is explained in §2.5 and different available bandwidth estimation approaches/tools are discussed in §2.6.

2.1 Emerging WiFi Networks

In this section, we first introduce multi-radio wireless mesh network as one of the emerging networks used to study coexistence interference characterisation and mitigation. Then, we explain how virtual access points work, while their effects on accuracy of location estimation is studied in Chapter 5. Finally, important features of high throughput WiFi networks based on the 802.11n standard are explained, while their effects on ABE are studied in Chapter 6.

2.1.1 Multi-radio Wireless Mesh Networks

Wireless mesh networks (WMNs) are defined as multi-hop static networks that provide connectivity – usually to the Internet – for clients who may be mobile or static. A WMN comprises of mesh routers connected together through their wireless interfaces in a mesh topology. A mesh router, as a main building block of WMNs, is usually equipped with an 802.11 interface – it is much cheaper in price compared to other wireless technologies such as UMTS and WiMAX equipment; and operates in unlicensed ISM frequency bands.

Several advantages of WMNs include being a cheap and fast technology in deployment; being a low cost in network operation (they do not have a high maintenance cost for wired networks such as cable networks and fibre optic networks); and being very low in energy consumption compared to the mobile networks. These are all promising deployment of WMNs to provide cheap and fast Internet connection services in rural areas and even in urban areas where fast deployments and low maintenance cost are important. Specifically, technologies like satellite, mobile data network and ADSL are either expensive (*e.g.*, cellular, WiMAX) or do not provide good quality connections to deliver traffic for new multimedia applications in rural areas [47].

Surprisingly, WMN deployments have received much attention in urban areas where alternative technologies (*e.g.*, ADSL, Cable TV, etc.) provide good quality Internet access with a reasonable price. Decreasing cost of portable devices, popularity of smart phones and Internet-based applications massively raise the demand for mobile data services. New technologies in mobile networks like HSDPA[1] and 4G LTE[2] have emerged to provide faster mobile data services. A cheaper, alternative solution for city-wide mobile Internet service is using public WiFi hotspots where WMNs are seen as a low cost to build and an operating backbone network for fast expanding coverage of the WiFi Internet service across a city [48].

However, WMNs like many other technologies are not scalable with ever-increasing bandwidth demands. To alleviate performance issues in WMNs, deploying multiple radios in a multi-radio platform has been introduced. Intuitively, adding more radios operating in orthogonal channels increases total capacity proportional to the number of radios; for example, adding another radio to a single radio platform would have doubled the communication capacity. However coexistence interference makes the benefit of having multiple-radio platforms marginal unless special care in the antennas place-

[1] High Speed Downlink Packet Access
[2] Long Term Evolution

ment, and assigned frequency channels are taken into account (this is further explained in §2.2.2). Taking into account the proposed recommendations in the multi-interface WMN system design (*e.g.*, minimum 40cm antennas separation) and configuration suggestions (*e.g.*, minimum three channel separation) results in a bulky systems and inefficient usage of the frequency channels. Our study in this thesis addresses some of these issues.

2.1.2 Virtual Access Points (VAPs)

VAP is a logical access point built by deploying a virtualisation technique to share a radio interface. The virtualisation approach can be a simple time sharing approach where the physical radio interface can be used only by one AP in each time slot. The main application of VAPs in the context of WiFi WLAN networks is supporting of different security policies for different classes of users.

It is now common that many WLANs used as business networks share some of their information services such as the Internet access with their guest users while access to many internal services are restricted for their employees. Figure 2.1 shows this deployment scenario. In this example, two VAPs (GuestWiFi, CorporateWiFi) are defined and their traffic is separated by different VLANs. Each VAP advertises itself every 100ms (the default interval value for the beacon transmission) which means the physical AP is transmitting beacons every 50ms.

However, VAPs advertise themselves as completely different APs, it is possible to detect the set of VAPs that belongs to a physical AP by looking at the last digits of the MAC addresses of the detected APs by WiFi scanning. Each VAP has its own virtual MAC address and it is advertised as BSSID of the WLAN network. BSSIDs of VAPs usually share the same first 40 bits of the physical AP's MAC address and only differs in the last two hexadecimal digits of the MAC addresses (*e.g.*, BSSIDs for GuestWiFi and CorporateWiFi in Figure 2.1).

2.1.3 High Throughput WiFi Networks

High throughput WiFi networks are first introduced in the 802.11n standard and enhanced in the 802.11ac standard. In this section, we review new features in both PHY and MAC layers that contribute to higher throughput. Note that it is not mandatory for all features specified in the 802.11n standard to be supported by 802.11n devices.

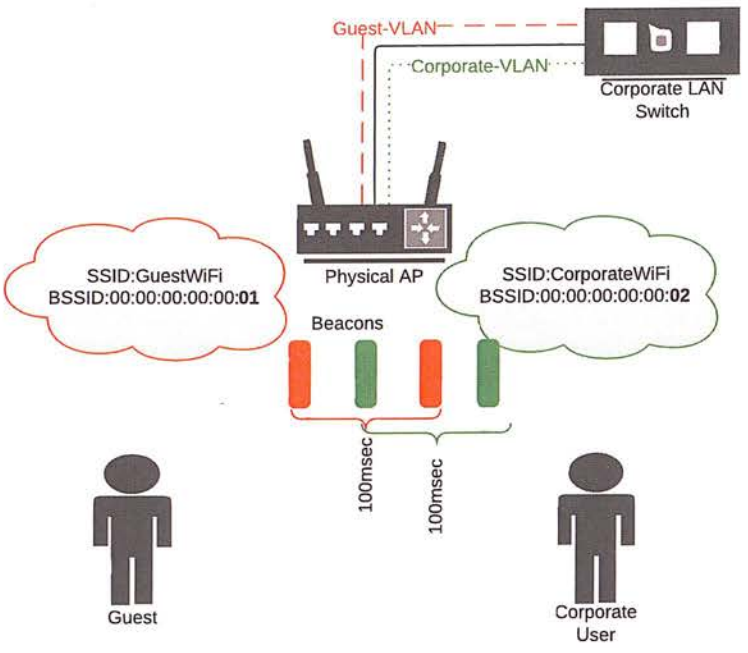


Figure 2.1: A deployment scenario for the VAPs application in a corporation.

2.1.3.1 PHY Layer Enhanced Features

Using spatial multiplexing, shorter guard intervals and increasing channel width are the main PHY enhancements.

Multiple Input Multiple Output(MIMO):

MIMO utilises multiple discrete antennas to transmit multiple data streams on the same channel. It increases communication range, reliability and link capacity, exploiting spatial diversity and spatial multiplexing. SDM (Spatial Division Multiplexing) is one of the MIMO features of 802.11n that increases the link capacity by concurrently transmitting different data streams over multiple streams, whereas, STBC (Space Time Block Coding) improves the transmission reliability by coding the same data using space-time coding and transmitting over different antennas. In the receiver the space-time decoding is deployed to combine the different received copies of a same data to extract as much information as possible [49].

Channel Bonding (CB):

Channel bonding as specified in the 802.11n standard combines two 20MHz channels to form a 40MHz channel [49]. Although it is introduced as an optional feature in the 802.11n standard, many WiFi manufacturers have implemented this feature in their

products. Enabling CB has been found to be harmful for transmissions in the 2.4GHz band as it increases interference from neighbouring channels [50] because of a limited number of orthogonal channels in this band.

Short Guard Interval (SGI):

Guard interval (GI) is a gap between each two symbols placed to mitigate inter-symbol interference due to delay spread of a channel. Supporting the SGI is optional in the 802.11n standard. By shortening the GI, the signal becomes more susceptible to longer delay spread channels. The standard OFDM data symbol for 802.11n has a 4 microsecond duration, comprising 0.8 microsecond for GI and 3.2 microsecond for data. This is the same as it is for legacy OFDM symbols, whereas with SGI, the guard interval is reduced to 0.4 microseconds which results in 3.6 microseconds overall for a whole symbol.

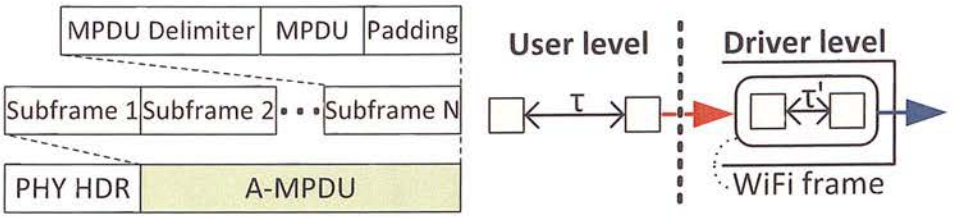
Frame Aggregation (FA):

Frame aggregation is a MAC layer enhancement and a key feature of 802.11n to improve protocol efficiency, i.e., to translate the physical layer bit-rates to comparable throughputs above the MAC layer. Protocol efficiency is a major concern underlying the design of 802.11n because of two reasons: (1) physical layer bit-rates are more than an order of magnitude higher compared to the earlier legacy standards of 802.11a/b/g (54Mbps vs. 600Mbps); (2) the protocol overhead (medium access, header overhead, inter-frame spaces) has a more harmful effect on higher layer throughput at higher physical layer rates. Without frame aggregation, the higher layer throughput achieved by 802.11n, with other major features enabled, is marginal [18].

The idea behind FA is simple: reduce the protocol overhead over several frames by transmitting them together. For FA to be effective, it needs a companion feature called block acknowledgement (similar to the selective ACK feature of TCP). 802.11n specifies two types of FA:

- Aggregate MAC Protocol Data Unit (A-MPDU)
- Aggregate MAC Service Data Unit (A-MSDU)

A-MPDU corresponds to aggregating multiple MPDUs (subframes) in the MAC layer, where MPDU (subframe) refers to a valid 802.11 MAC frame with MAC header, one IP packet as payload and a frame check sequence (FCS). A-MSDU, on the other hand, aggregates several IP packets above the MAC layer and puts them into one MAC frame with a common MAC header and FCS. A-MPDU is the most widely supported and



(a) A-MPDU type (b) System level view

Figure 2.2: Illustration of frame aggregation.

popular option for several reasons: (1) The A-MPDU option is typically implemented in the firmware and requires less host software support within the device driver; (2) it allows for greater reliability because each subframe has its own FCS and aggregated frame transmission can benefit from block acknowledgements; (3) it allows greater level of aggregation with A-MPDU size up to 64KB, whereas A-MSDU size is limited to around 8KB; (4) it is also flexible in that it can support multiple QoS traffic classes, one per subframe. Figure 2.2(a) illustrates frames aggregation in an AMPDU frame.

2.2 Interference

Interference is an unwanted disturbance like noise that hurts the communication performance. It results from concurrent transmissions from other wireless devices. Those devices could belong to the same network or nearby networks, or could be using different communication technologies such as WiFi and ZigBee devices, both working in the 2.4GHz band [51]. It could also be that interfering devices are within the same host/router, as would be the case with multi-radio platforms. Two types of interference studied in this thesis are: *Co-Channel Interference (CCI)* and *coexistence interference*. CCI happens when two or more devices share the same channel while coexistence interference happens when two or more radio interfaces are working in very close proximity (less than three meters) or placed together to make a multi-radio platform. In the following, further explanations are provided.

2.2.1 Co-channel Interference

The CSMA/CA[3] protocol is proposed to avoid collisions due to concurrent access to the channel. It comprises of two functional modules: *Carrier Sensing* and *Collision Avoidance*. Carrier sensing is responsible for sensing whether another node is transmitting before starting a transmission. For collision avoidance, stations use a random back-off time after each frame transmission or optionally transmit Request To Send (RTS) and receive Clear To Send (CTS) control frames to reserve the channel prior to transmitting [52]. Co-channel interference is a result of failing or misbehaving in channel access mechanisms. For example, in the hidden node problem [53], co-channel interference is happening because one station cannot hear/sense that the other station is transmitting owing to the distance between the two stations which are out of the communication range of each other; sometimes even close stations might not be able to hear each other due to high radio signal distortion because of reflection, diffraction and scattering.

Collision avoidance mechanisms proposed in CSMA/CA are also shown not scalable by increasing number of stations and APs. For example, [54] using Orbit testbed to systematically investigate the effect of access points and stations density on the performance of different applications such as web and multimedia. They show that increasing the number of APs can significantly increase the presence of co-channel interference; increasing the number of APs to four with 75 stations results in fall in cumulative TCP throughput (8.7 Mbps vs. 127.88Mbps for the single AP scenario).

2.2.2 Multi-radio Coexistence Interference

Three reasons are identified as causes for poor performance of multi-radio platforms when radios are working in different bands/channels. They are *transmitter noise*, *receiver blocking* and *inter-modulation* and are collectively referred to as *coexistence interference* [15]. Note that coexistence interference is a more general term than *adjacent channel interference (ACI)* which is usually referred to in the literature. In the following, definitions for each underlying reason are provided.

2.2.2.1 Definitions

Transmitter Noise: Transmitter noise or out-of-band emission, which causes ACI, occurs due to imperfect filtering applied by the bandpass filter or transmit spectrum

[3]Carrier sense multiple access with collision avoidance

mask while transmitting signals.

ACI between orthogonal channels happens when two interfaces are close together or coexist in a platform. Two channels are adjacent if they have different central frequencies in a band. They could be two orthogonal channels such as channels 36 and 40 in the 5GHz band or channel 1 and 6 in the 2.4GHz band. Figure 2.3 illustrates how concurrent transmission on adjacent channels can act as a high power noise for transmission on the other interfaces set up in different orthogonal channels. The rectangular shape in the figure shows an ideal transmit spectrum mask, which is usually assumed in simulation studies, and does not exist in reality.

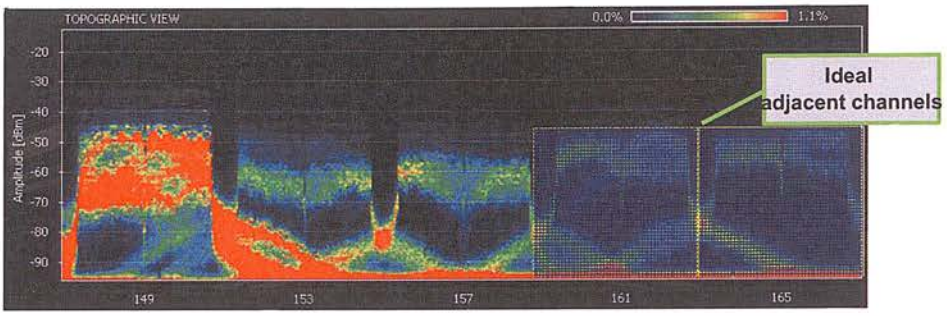


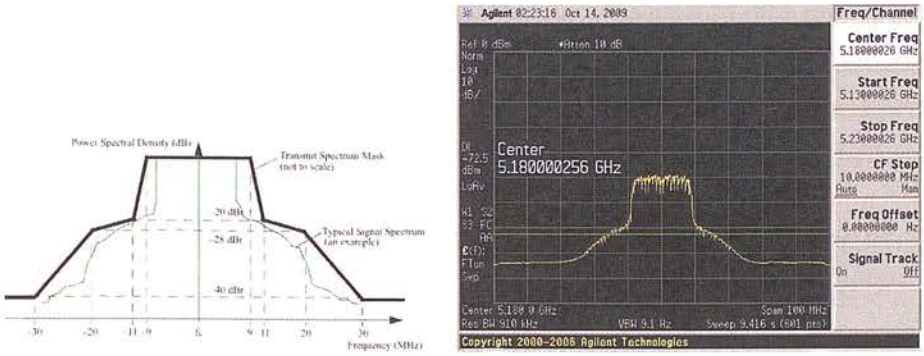
Figure 2.3: ACI observed in the 5GHz band used 802.11a devices.

The IEEE802.11 standard gives multiple definitions for the transmit spectrum masks. It has different specifications for transmissions with 20MHz channel width and transmissions with 22MHz channel width – as is the case for 802.11bg. For example, the transmission spectrum mask for operation in 20MHz channels is specified as: *"the transmitted spectrum shall have a 0 dBr (decibel relative to the maximum spectral density of the signal) bandwidth not exceeding 18MHz, 20dBr at 11MHz frequency offset, 28dBr at 20MHz frequency offset, and 40dBr at 30MHz frequency offset and above. The transmitted spectral density of the transmitted signal shall fall within the spectral mask, as shown in Figure 2.4(a)."*

Figure 2.4(b) shows an example of an output signal that is observed in our lab from a Ubiquiti XR5 WiFi card measured by deploying a spectrum analyser.

Inter-modulation and Receiver Blocking: Two types of Inter-modulation interference are recognised based on where they are generated: *transmitter inter-modulation* and *receiver inter-modulation.*

Inter-modulation generated in the transmitter usually happens in the non-linear power amplifier of the transmit circuits of a radio interface. It happens due to closeness



(a) (b)

Figure 2.4: (a) Transmit spectrum mask specified in IEEE 802.11 standard [3]; (b) Output spectrum observed for Ubiquiti XR5 WiFi card transmitting in channel 36 of the 5GHz frequency band.

of two transmitting antennas which causes the signals of one transmitter to be received through the antenna of the other transmitter and coupled with its transmitting signals in the final power amplifier. The mixing of two signals produces undesired signals at different frequency intervals away from the carrier signal [55].

Inter-modulation can also cause the receiver to be desensitized. It happens by receiving strong inter-modulate signal (even in different frequency of desired receiving signal) in the automatic gain control of the receiver RF circuit. It makes the receiver reduce the receiving gain which also reduces the receiving signal power of the desired signal to a point that cannot be demodulated, causing receiver blocking. In other words, receiver blocking is generally due to limited dynamic range of power amplifier and A/D converter in the receiver. If total input power at the receiver is more than the blocking limit (*e.g.*, -30dBm at 2GHz for WiFi) the received signal strength degrades.

Both inter-modulation and receiver blocking can results from different system design issues including: Board crosstalk, WiFi cards radiation leakage, inadequate separation between antennas [13].

2.2.2.2 Multi-Radio Coexistence Interference Characterisation

A common measurement approach in characterising coexistence interference is measuring its effects on traffic performance (*i.e.*, UDP/TCP throughput) usually in a three-node topology with a multi-radio middle node. In many characterising studies, multi-radio platforms are not commodity hardware usually used as routers or APs in real

deployments of 802.11-based wireless networks; instead, collocated laptops sitting together (*e.g.*, [14]) or workstations equipped with multiple WiFi PCI cards (*e.g.*, [13]). In testing scenarios, effects of the following different parameters on traffic performance are studied.

- *MAC/PHY parameters:* frequency channel, transmission power and transmission coding rate and modulation.
- *System aspect parameters:* distance between antenna, shielding of radio circuits.
- *Traffic patterns toward the multi-radio node:* transmitting on all radio interfaces (TX-TX), receiving on all radio interfaces (RX-RX) and a combination of concurrent receiving or transmitting for each interface (RXTX).

It is worth mentioning that the listed parameters are not completely independent parameters; for example, the amount of transmission power, distance between antennas and separation between frequency channels influence on each other — shorter channel separation needs more antenna separation and lower transmission power.

As explained before, receiver blocking can results from WiFi signal radiation leakage and inadequate separation between antennas [13]. Radio signals can leakage from radio chipsets, connectors on the WiFi card and antennas' pigtails. To mitigate the radiation leakage, shielding with metal plates is the main proposed solution in the literature [13, 56]. In the following, we will discuss more about the impact of antenna placement (antenna separation).

2.2.2.3 Impact of Antennas' Placement

The work of Robinson et al. [13] is one of the first works that considers system issues with multi-radio systems including distance between antennas. They use a two-hop network topology comprising of three workstations equipped with 802.11b-based wireless PCI cards and measure throughput for characterising coexistence interference. Their findings recommend having **one metre** antenna separation to remedy the effect of coexistence interference for radio interfaces set up in orthogonal channels.

However, the suggested antenna separation distances are different in separate studies — the minimum antenna distance reported by [14] is at least **38 cm**. They use two laptops equipped with 802.11b/g cards, sitting together and connecting through their Ethernet ports to mimic a multi-radio platform. A recent study that characterises interference in multi-radio platforms based on the new 802.11n standard recommends **one**

meter separation between antennas [57]. Authors of [30] also study the effects of antenna separation for operating in the 5GHz band using commodity network platforms usually used in real deployments. They report different required antenna distances for different modulation schemes (OFDM vs. DSSS) and different bands (2.4GHz vs. 5GHz). For relay node scenario (RX-TX), a minimum of **40 cm** antenna separation is recommended when transmitting either on 2.4GHz using DSSS modulation (802.11b) or on 5GHz using OFDM modulation (802.11a). Surprisingly, they report at least **160 cm** separation between antennas is required for using OFDM modulation in 2.4GHz band. [14] have also reported the benefit of DSSS modulation over OFDM modulation.

Based on the mentioned studies, a minimum of 40cm antenna separation is required considering enough separation between channels and proper transmission power level which varies in different setups. Alternatively, shielding between antennas (*e.g.*, [57]), using directional antennas instead of omnidirectional antenna even for indoor deployments (*e.g.*, [57, 58]) and using external bandpass radio filters (*e.g.*, [28]) are proposed in the literature to shrink the distance between antennas and consequently dimensions of multi-radio platforms. Specifically, external bandpass filters (in addition to those built in the wireless card) are introduced to further attenuate the unwanted adjacent channel signal power. The majority of bandpass filters, however, are designed to work in a specific frequency range, which limits their use across a range of channels. This limits the operation in multiple bands and the flexibility in choosing between frequency channels. External bandpass filters are usually bulky and expensive. Using directional antenna or shielding between antennas would also affect the signal propagation compare to omnidirectional propagation.

Natchtigal *et al.* in [30] report that even with 320 cm antenna separation configuring two radio interfaces in the direct adjacent channel (*e.g.*, channel 36 and channel 40 from 5GHz band) is not possible. This, along with other studies shows the importance of channel separation.

2.2.2.4 Impact of Frequency Channels

As explained in §2.2.2.1, ACI happens because of out-of-band emission owing to imperfect filtering applied by the bandpass filter or transmit spectrum mask. To lessen the harmful effects of receiving high power from transmission on other radio interfaces, it is required to further separate the antenna, reduce the transmission power (*e.g.*, to as low as 5dBm as reported by [14]) or increase the distance between channels in the

frequency domain which is called *channel separation*.

Studies in the literature show that antenna and channel separation should be considered together to remedy the effect of interference. For a minimum of 40cm antenna separation, a minimum of three orthogonal channel separation is required when operating in 5GHz band (*e.g.*, if channel 36 is assigned to one interface the other interface channel should set to channel 52 or higher channels) [30, 29]. [30], concerning transmission on the 2.4GHz band using OFDM did not find any two channels to prevent throughput degradation for the same antenna distance of 40cm. [14] also study the effect of channel separation and transmission power and show that increasing channel separation helps transmitting with higher transmission power consequently improves transmission range.

2.2.2.5 Impact of Traffic Direction

Three possible application layer traffic patterns are studied in the literature as different behaviours have been observed. They include:

- *TX-TX*: All interfaces are transmitting simultaneously to the other nodes.
- *RX-RX*: All interfaces are receiving traffic from other nodes.
- *RX-TX*: Some interfaces are transmitting to the other nodes and some of them are receiving from other nodes.

However, all these traffic patterns in the MAC layer fall in to the RX-TX pattern as, for every successful reception of a unicast frame, an ACK frame will be transmitted back; so, at a MAC layer we do not have directional unicast traffic. Note that ACK frames are very small and transmitted with the lowest coding rate to increase the chance of successful delivery and are therefore more robust against interference.

[13, 14, 30] report different behaviours from each of the three traffic scenarios where RX-TX is more affected by the interference as is explained in the following.

[14] explains the following two effects happening for different traffic patterns:

- **Spurious carrier sensing:** Interfaces operating in the adjacent channel with packets in their transmission queues defer from transmission because they mistakenly sense that the channel is used by other co-channel nodes. However, it is due to out of band emission power from adjacent channels (this phenomena is directly observed by using the measurement technique introduced in §4.3.3).

- **Interference noise:** Transmission on other interfaces acts as a high power noise for receiving interfaces configured in adjacent channels.

Hence, in TX-TX spurious carrier sensing causes deferral of transmission that affects the throughput, while for RX-TX and RX-RX traffic patterns, interference acting as noise causes packet loss on traffic of the receiving interfaces.

Effects of traffic patterns on required channel separation. [30] study effects of antenna distance and channel separation for different traffic pattern scenarios. They show that for the RX-RX scenario the minimum separation is two channels while it is three channels for TX-TX and four channels in the RX-TX scenario when antenna separation distance is 40 cm and all interfaces are operating in channels from 5GHz band.

In summary, all of the reviewed studies report that multi-radio coexistence interference can adversely affect traffic performance. To mitigate this interference, antenna separation of at least 40 cm and a minimum of three orthogonal channel separations are proposed. Large antenna distance results in bulky platforms which is not suitable to be deployed in indoor environments. Most of these studies rely on packet delivery ratio measurements to make inferences about different underlying reasons in MAC and PHY layer. In Chapter 4, we propose using different antenna polarisation (§4.4) to remedy the effects of coexistence interference and having more compact multi-radio platforms, and propose using a measurement technique to have a direct view of presence of ACI from the MAC layer.

2.3 Antenna Polarisation

The antenna is an important element of any wireless device — it converts electric current into electromagnetic wave and vice versa when receiving an electromagnetic wave. As it is shown in Figure 2.5, the electrical (E) and electromagnetic (H) fields are perpendicular to each other and to the direction of the wave propagation [59].

The interplay between electric and magnetic fields store energy and hence carries power along the propagation direction (Pointing vector). Information to be transferred is embedded in the wave by variation, modulation, in parameters of the wave (amplitude, frequency or phase).

Any electromagnetic wave is characterised by several fundamental properties, including its frequency, speed, amplitude, phase angle, polarisation, direction of propa-

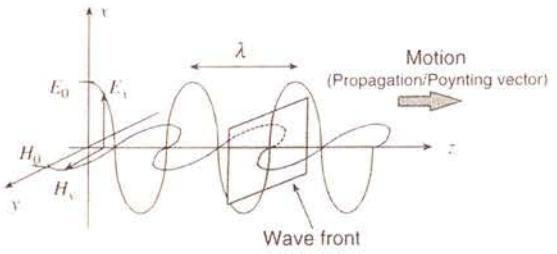


Figure 2.5: A plane wave propagating through space at a single moment in time [59].

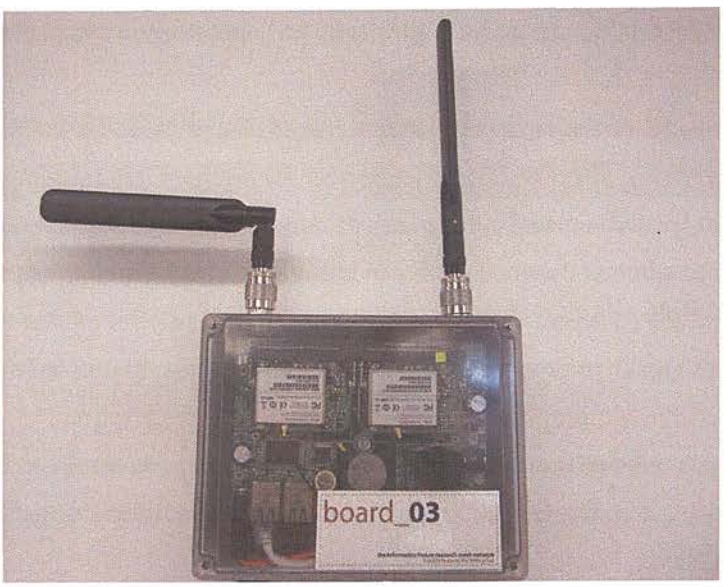


Figure 2.6: Omnidirectional antenna in different polarisation in a multi-radio mesh router.

gation. Polarisation is defined in [59] as "the alignment of the electric field vector of a plane wave relative to the direction of propagation". If the wave is generated by a vertical wire antenna compared to the ground, then the wave is said to be vertically polarised; a wire antenna parallel to the ground primarily generates waves that are horizontally polarised. Figure 2.6 shows a dual-radio router equipped with a vertical polarised antenna (right antenna) and a horizontally polarised antenna (left antenna). Later in Chapter 4, polarisation diversity (using antennas in different polarisations) is exploited to remedy the effects of coexistence interference.

2.4 Urban WiFi Monitoring Systems

There are two main approaches employed for urban WiFi characterisation: *deploying networks of WiFi sensors which we refer to as fixed infrastructure* and *wardriving* which is a mobile approach. This section discusses the advantages and disadvantages of each approach and gives some examples of real deployments of them. Then, the successfulness of the *Crowdsensing* approach in other applications and its potentials for WiFi networks characterisation are discussed.

2.4.1 Fixed Infrastructure

In a fixed infrastructure approach, a set of monitoring devices are positioned across the area of interest. Argos [22] is an urban WiFi monitoring system that exemplifies this approach. It is based on a deployment of a stationary set of 2.4GHz sensors (sniffers); these sniffers are interconnected as a wireless mesh network operating on a separate 900MHz channel. The contribution of Argos lies in efficient mechanisms for coordinated channel sampling by multiple sensors and a collection of monitoring traffic, both aimed to cope with the limited backhaul mesh capacity.

The study reported in [5] presents another example following this approach; here measurement data is manually retrieved from monitoring equipment. A laptop with a GPS receiver and two USB dongles, AirPcap Nx [60] and WiSpy DBx [61], is deployed for a day at some chosen locations spanning different WiFi environments (houses, apartments, cafés and shopping centres).

From a characterisation and monitoring perspective, the requirement to deploy a dedicated infrastructure makes this approach expensive, especially for fine-grained spatio-temporal mapping. As an intermediate approach, existing mesh or vehicular networks are exploited for monitoring purposes. For example, Hare et.al in [62] introduce the WiRover framework which provides vehicular open Internet wireless access for passengers on buses. The framework is used as a monitoring infrastructure to study the Internet usage behaviour of bus passengers. As another example, the free access Google WiFi network in Mountain View is exploited in [63] to study different classes of users based on the devices used for the Internet connection (*i.e.*, laptops, fixed location access devices, and smartphones). The authors also show how different classes of users are geographically distributed between residential, commercial and transportation areas of the city.

2.4.2 Wardriving

Wardriving has been the most common approach taken for urban WiFi characterisation. It typically involves a group of wardrivers, each carrying a specialised laptop-class WiFi and a GPS-equipped device running wardriving software (*e.g.*, inSSIDer [64]), possibly with a custom antenna, going around the city to locate existing WiFi APs. During this operation, the wardriving software is often the only application running on the device [23]. The approach underlying the study reported in [4] can also be seen as a wardriving approach where measurements are collected while walking around selected London neighbourhoods. There are some public databases like WiGLE [65] that aggregate data from wardriving campaigns.

Typical use of wardriving data and the resultant mapping of WiFi APs is for localisation (*e.g.*, Skyhook [66], Place Lab [67]), as a more reliable, faster and energy-efficient alternative to using GPS. The authors of [27] report another use for wardriving that is aimed at assessing the feasibility of vehicular Internet access via open WiFi APs; their companion website [68] shows a map of APs found from the wardriving exercise in the Boston area. Like typical wardriving studies, [27] also makes use of a custom hardware/software platform.

In summary, as the authors in [69] note, *wardriving is an expensive and tedious operation. As such it may be impractical for fine-grained and continual WiFi monitoring.*

2.4.3 Mobile Crowdsourcing and Crowdsensing

Mobile crowdsensing, according to [70], is defined as *"individuals with sensing and computing devices collectively share data and extract information to measure and map phenomena of common interest"*. It bears similarity to wardriving but eases the burden on the participants and makes use of off-the-shelf smartphones with measurement software running in the background. Thus it has the potential to enable *cost-effective*, *fine-grained* and *continual monitoring*.

Deployments of mobile crowdsensing in other applications. The ubiquitous nature of mobile phones, rapid migration towards smartphones and the presence of several built-in sensors on smartphones on the one hand and communication interfaces to share collected information (*e.g.*, in the cloud) on the other hand have led to a significant interest in mobile phone sensing.

Given different embedded sensors of smartphones such as accelerometer, digital

compass, gyroscope, barometer, GPS, microphone, proximity sensor, and camera, a variety of mobile sensing applications in different domains have seen proposed in recent years, such as air pollution monitoring [71, 72, 73] using hand held air pollution sensor or urban noise monitoring in NoiseSPY [74] turning mobile phone into a noise sensor using its microphone.

Crowdsourcing-based mobile network measurements. In the context of mobile cellular network measurement, *crowdsourcing* approaches have received considerable attention. The problem of useful yet scalable crowdsourcing-based mobile network measurement is tackled in [75, 76]. For example, WiScape ([76]) is a crowdsource-based framework for characterisation of connection performance for wireless networks using portable devices like mobile phones or laptops with mobile wireless interfaces [76, 77]. Similar to WiScape, Speedtest.net is a broadband connection analysis framework to measure the Internet speed and latency against distributed servers. It collects measurement data from different sources including its website and its mobile phone application. OpenSignal [78] and Mobiperf [79] also represent passive and active crowdsourced mobile network measurement systems with freely available mobile apps.

Crowdsensing for WiFi characterisation. In the context of WiFi network characterisation, [80] uses mobile crowdsourced datasets to report analysis and comparison of mobile Internet access performance between WiFi and cellular connections using *speedtest* mobile app based active performance measurements (download/upload speeds and latency). In contrast to [80], our objective is to use mobile crowdsensing for characterisation of urban WiFi deployments. Pazl [81] is a recent work which aims at WiFi monitoring within indoor environments and addresses the associated challenge of locating measurements via a hybrid localisation mechanism that combines pedestrian dead reckoning with WiFi fingerprinting. In [82], the authors propose a system to detect and track WiFi enabled smartphones using off-the-shelf access points as monitoring stations, the inverse of the problem that we consider in Chapter 3, which is to detect the presence of WiFi networks using commodity smartphones.

2.5 WiFi Fingerprinting for Localisation

WiFi fingerprinting is one of three main techniques for localisation which work based on the receiving signal for APs in the network. The strongest base station method is

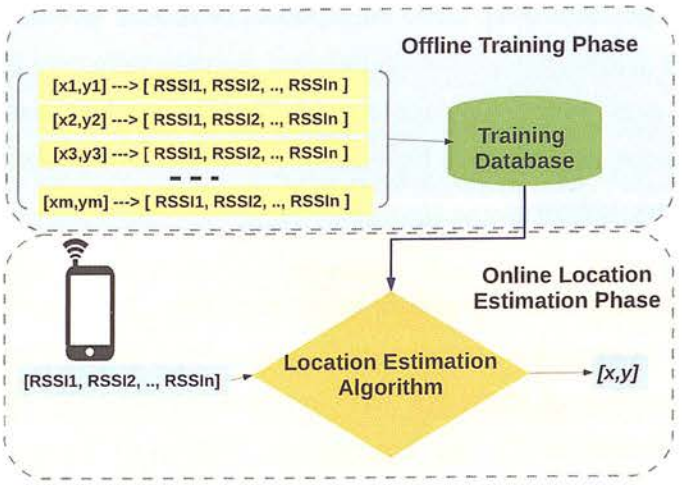


Figure 2.7: Building blocks of a typical WiFi fingerprint system.

computationally the simplest method. It is as simple as choosing the position of the AP with the strongest received signal as the location of the user. This technique does not have enough precision mainly due to the complexity of signal propagation in indoor environments. The other technique is using radio propagation modelling. It takes the signal path loss information calculated based on theoretical models such as free space path loss to estimate the distance to the observed APs, then the location of the user is calculated by using a geometry technique such as trilateration and triangulation. The accuracy of the location estimation is very dependent on the accuracy of the radio propagation model, which can be quite challenging in multipath rich indoor environments.

To prevent having to go through the process of radio propagation modelling, fingerprinting techniques are proposed and are becoming very popular for indoor localisation owing to its better precision in location estimation as compared to other WiFi based localisation schemes. Early WiFi fingerprinting systems including RADAR [33] and Horus [83] rely on an initial training phase to construct the fingerprint database for use as a reference in the positioning phase later; but the training phase can be quite time consuming and expensive. More recent WiFi fingerprinting systems make this training phase automated via crowdsourcing using various mechanisms with increasing sophistication (*e.g.*, Redpin [84], OIL [85], Zee [86]).

A WiFi fingerprinting system normally consists of the following two phases:

- **Site survey/ offline/ training phase.** The first phase involving building a fin-

gerprint database or constructing a radio map through measurements associated with true locations.

- **Online/ runtime/ positioning/ tracking phase.** The second phase involving estimating the absolute location by looking up the closest matching between samples in the radio map database and those online collected by the user's device.

Figure 2.7 shows the building blocks of a two phase WiFi fingerprinting system.

Offline phase. As is illustrated in the offline phase, signal strength of observed APs in each location is measured and then stored in the fingerprint database (note that usually multiple samples are collected for observed APs). However, fingerprint definition is not limited to the received signal strength. In §5.2.3, we study performance of five fingerprint definitions across different environments. Three fingerprints are defined based on the RSSI — signal strength; stability of signal strength indicator which is calculated as the standard deviation of RSSI; and RSSI variation among different measurement locations. The other group of fingerprint definitions is defined based on the visibility of APs. In indoor environments — because of fading and multipath propagation of signals — it is very likely that different sets of APs are seen in each location. It even happens between different measurements in a location because of either signal fading or losing beacons due to co-channel interference. Hybrid fingerprint definitions are presented by combining definitions from the two mentioned groups [87]. For example, constance visibility of an AP across multiple WiFi scannings and maximum signal strength of receiving beacons from different APs are combined to define a new fingerprint definition as: APs that are more constantly visible are chosen first, then among them those that have a higher signal strength are finally chosen (more explanations can be found in §5.2.3).

There are studies comparing different WiFi fingerprinting techniques (*e.g.*, [88]) and analysing the properties of WiFi signals as they pertain to location fingerprinting (see [89] and references therein). A number of factors are now recognised to have an impact on the accuracy of WiFi fingerprinting systems to varying degrees, including user orientation, temporal and spatial variations of WiFi signals, device hardware, transmit power, and number of measurement samples [33, 83, 89].

Online phase. In the online phase, different location estimation techniques are proposed in the literature: *deterministic techniques* or nearest neighbour (NN) techniques and *probabilistic techniques*.

In NN techniques, the distance between two RSSI vectors from the training data

set and the on-line data set is calculated. Then, the user's location is determined by finding a cell that has the closest distance to the online reading vector [90]. Different NN techniques proposed in the literature differ in calculating the distance between vectors. RADAR [33] is one of the first schemes that uses WiFi fingerprinting with a NN localisation algorithm.

Horus in [91] proposes a probabilistic approach supposing RSSI temporal distribution follows the Normal or Gaussian distribution. Although this assumption is questioned by other studies (*e.g.*, [89]), Horus is widely considered as one of the most practical WLAN fingerprinting systems and is often used as a benchmark for comparative studies. More explanation about the localisation algorithms is given in §5.2.4.

2.6 Overview of Available Bandwidth Estimation Approaches

In this section, some of the common metrics and concepts — usually referred to in available bandwidth estimation studies — are introduced. Tools based on differences in their measurement techniques are also listed later in this section and more explanations are given for tools which are under study in Chapter 6.

2.6.1 Terminology

The following definitions explain different terms and metrics in the context of available bandwidth estimation.

Cross traffic. Cross traffic is traffic that shares the link capacity with measurement traffic.

Capacity and effective capacity. Capacity is a maximum possible bandwidth a link or path can deliver (*i.e.*, in Figure 2.8, capacity of the path is defined as the minimum capacity of the links in the traffic path). Capacity of an 802.11-based link is defined as the physical link transmission rate which is dependent on the chosen coding and modulation schemes for transmission of a frame and it can be different between transmissions of different frames. Effective capacity is defined in [92] as the amount of bandwidth effectively available considering protocol overheads such as channel accessing time and MAC layer acknowledgement in 802.11-based communications.

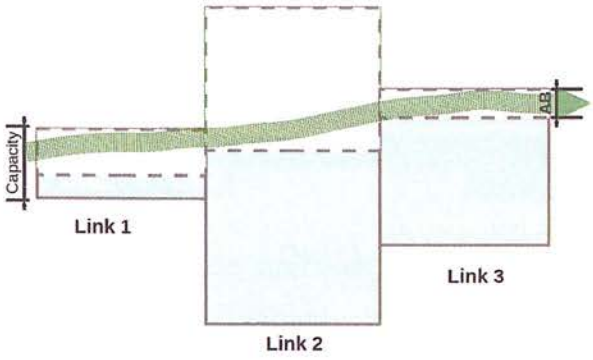


Figure 2.8: Capacity and Available Bandwidth of a path with three different link capacities and available bandwidths.

Available bandwidth. Available bandwidth (AB) is defined as maximum spare bandwidth of a link or an end-to-end path (AB for a path is defined as the minimum ABs for links in the traffic path, see Figure 2.8). While capacity is more affected by the underlying technologies, available bandwidth is more influenced by the cross traffic load making it a more dynamic metric. [93] defines the protocol-dependent bandwidth as a sustainable rate the application can achieve which depends on the application's behaviour.

Narrow link and tight link. Narrow link refers to the link in the path which has the lowest capacity (*e.g.*, Link1 in Figure 2.8), while the tight link refers to the link that has the smallest available bandwidth and it might be a different link in the path (*e.g.*, Link3 in Figure 2.8).

Probe. Probes are packets transmitted for the purpose of available bandwidth or capacity measurement.

Probing techniques. Probing techniques are specified based on the pattern of transmitting probes. In what follows, some probing approaches are explained:

- **Single packet probing:** sending a packet as a probe. This approach is usually deployed when sender and receiver of the probes are synced and one-way de-

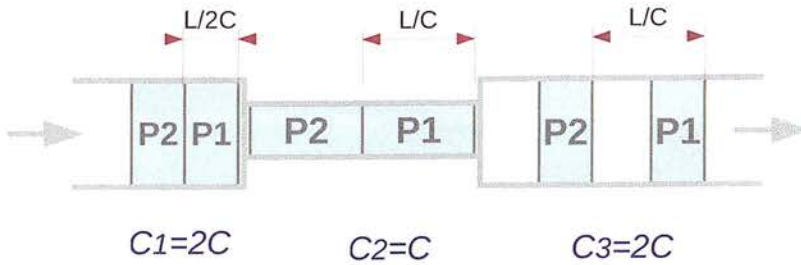


Figure 2.9: Packet pair probing.

lay transmission time can be calculated using differences between receiving and sending time.

- **Packet-pair probing:** sending a pair of probes back-to-back. Figure 2.9 shows a pair of probes (P1 and P2) traversing a path with one tight link (the middle link) with capacity half of the two other links in the path (C2=C). Measuring the time between the two probes at the receiver and having the size of the probe(L) gives the capacity of the path. Packet pair approaches are usually exploited for measuring the link capacity as sending probes in pairs reduces the chance of being interleaved with the cross traffic.

- **Variable packet size probing (VPS):** sending different sized back-to-back probes; it is also known as tailgating probing [94].

- **Packet train probing:** sending multiple probes that are separated based on the transmission probe rate (they are not sent back-to-back as in the packet-pair probing, see Figure 2.10). Packet train probing is usually exploited for available bandwidth estimation rather than link capacity estimation as cross traffic has chances to interleave between the probes and affect the probes interval time (T).

- **Multi-layer probing:** sending probes from different layers of the protocol stack. This probing approach is deployed to measure more refined ABE in the application layer.

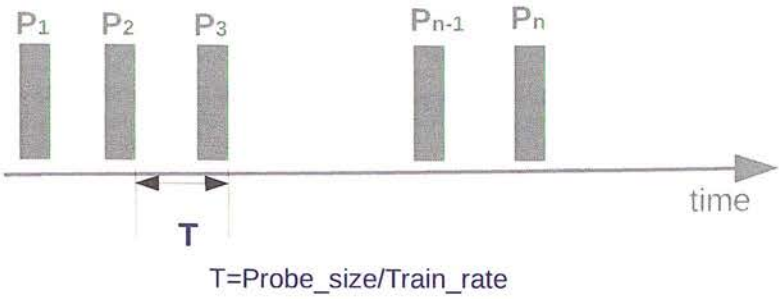


Figure 2.10: Packet train probing.

Dispersion time. Dispersion time is defined as separation time between two probes transmitted together owing to passing through a narrow link or the existence of cross traffic (In Figure 2.9, dispersion time is L/C). In the context of WiFi networks, dispersion time is also affected by the contending traffic and channel quality.

Self-induced congestion. Packets' inter-space gets increased if transmission rate is more than the available bandwidth causing packets to get queued in the tight link and increasing the delay.

Rate response curve. It shows the relation between the input rate and output rate when using self-induced congestion approaches.

Inference techniques. Inference techniques are deployed to infer the available bandwidth or capacity from the probes' measurement. Two main active ABE inference techniques exploited in different tools are Probe Gap Model (PGM) and Probe Rate Model (PRM) (other techniques such as One Way Delay (OWD) are only exploited by a few tools like ProbeGap [38]).

- **Probe Gap Model (PGM):** this model relies on the observation that successive probe packets traversing a tight link undergo increased dispersion in time due to cross-traffic. The rate of cross traffic (alternatively, the utilisation of the tight link) is estimated via an increase in dispersion experienced by back-to-back probe packets. Available bandwidth is then computed using the estimated cross-traffic rate and capacity; the latter is either assumed to be known or separately estimated. Examples of PGM based tools include: Spruce [95] and WBest [40]. For example, WBest start sending 30 packet pairs to estimate the capacity of

the link and then sending packet train of 30 probes in the rate of the estimated capacity to calculate the average dispersion time between probes. The available bandwidth is calculated using the fluid model supposing the capacity and the average dispersion time are known. In fluid model it is supposed that traffic can be modelled as FIFO queue and probing traffic has the same share of the traffic before and after the queue [40].

- **Probe Rate Model (PRM):** This model is based on the notion of self-induced congestion. The essential idea is that when the probe traffic rate exceeds the available bandwidth of the path, the measured reception rate of probes starts lagging behind their sending rate (or equivalently, the inter-probe arrival times measured at receiver keep increasing). The maximum probe traffic rate at which this transition occurs is taken as the available bandwidth estimate. PRM based tools are usually iterative, spanning several rounds to probe at different rates. Several tools belong to this category, including Pathload [96], TOPP [93], DietTopp [39] and pathChirp [46]. These tools differ in their probing traffic patterns and receiver-side statistical analysis mechanisms.

- **One Way Delay (OWD):** OWD-based tools compute available bandwidth using the link capacity and the fraction of time the link is observed to be idle . The percent of time link is considered to be idle is measured by looking for considerable increasing in delay of receiving probes.

Bandwidth estimation approaches are divided — based on their **intrusiveness** to the network users traffic — to *passive* and *active* measurements. Passive measurement techniques are usually based on monitoring/observing network traffic and other network statistics to estimate the AB while active approaches send probes in the network and study changes happened to probing traffic at the receiver, which makes them intrusive to the existing network traffic. The effect of the probing traffic is more intrusive in WiFi network context considering contention in accessing the channel and the MAC layer acknowledgement.

2.6.2 Passive Available Bandwidth Estimation Tools

Passive measurement techniques are based on monitoring a link over a period of time, measuring idle and busy time, and calculating the available bandwidth as a percent

of time the link is sensed being idle. The main advantage of passive measurement approaches are being non-intrusive compared to the active traffic.

Available Bandwidth Estimation(ABE) [97] and Improved Available Bandwidth estimation (IAB) [98] are among previously proposed passive available bandwidth estimation approaches for wireless networks. cPEAB [99] proposed to enhance previous tools by considering MAC layer control traffic and the effect of the hidden and exposed nodes. The available bandwidth is estimated by measuring waiting and back off delay, packet collision probability, acknowledgement delay and channel idle time supposing that hidden/exposed nodes' information are given by the network management system.

The main problems with passive approaches that makes them less popular are: i) their high dependency on information from the system and network device driver ii) requirement of having administrative access for capturing packets and iii) unfitness for End-to-End ABE (e.g., between a server in the Internet and a laptop connected through home broadband).

2.6.3 Active Available Bandwidth Estimation Tools

Active measurement techniques are based on sending probes with different rates, sizes, numbers or patterns and measuring delays or loss to estimate the available bandwidth. They may differ in probing techniques (*e.g.*, packet pair, packet train, packet chirp) as well as estimation/inference techniques (*e.g.*, PGM vs. PRM). Table 2.1 lists some of the ABE tools introduced; they are categorised based on their inference techniques, inference metrics and estimation metrics (available bandwidth or link capacity). Note that the inference metric is an underlying metric used by the inference techniques to estimate available bandwidth. For example, the rate of the train is exploited by many PRM tools, and the dispersion time is usually used by the PGM tools. Sometimes the combination of both metrics are deployed as is for IGI/PTR [100].

As it is illustrated in Table 2.1, many proposed tools work based on the PRM technique. While these tools differ in the inference metrics and filtering approaches which are used to choose between different ABE measurement values (*e.g.*, mean or median). Generally, they work in a similar fashion in that they narrow down the possible values of the available bandwidth by increasing probe sending rates and identifying a transitioning point where the sending rate becomes larger than the receiving rate. One notable difference between PRM and PGM tools is that while other techniques are required to measure or know bottleneck link capacity somehow, PRM-based tools have

Table 2.1: Available Bandwidth Estimation Tools.

Tools	Inference Tech.	Inference Metric	Estimation Metric
Pathrate [101]	PGM	dispersion time	Capacity
CapProbe [102]	PGM	delay, dispersion time	Capacity
Spruce [95]	PGM	dispersion time	AB
Wbest [40]	PGM	dispersion time	Capacity,AB
DietTOPP [39]	PRM	received train rate	AB
Pathload [103]	PRM	received train rate	AB
Yaz [104]	PRM	received train rate	AB
BART [105]	PRM	received train rate	AB
SLDRT [106]	PRM	chirp rate	AB
IGI/PTR [100]	PRM	dispersion time, received train rate	AB
PathChirp [46]	PRM	chirp rate	AB
ProbeGap [38]	OWD	one way delay	AB

no such constraint.

The majority of tools and techniques proposed in the literature are designed to be deployed in wired data networks. Among the listed tools in Table 2.1 — for the study in Chapter 6 — three tools are chosen: **WBest**, **DietTopp** and **pathChirp**. They employ the two main ABE models (PGM and PRM) described above. They are also among the tools that are also known to work in the WLAN environments. WBest has been specifically designed for (legacy) 802.11 wireless LANs and represents the PGM class of ABE tools. DietTopp, on the other hand, falls in the PRM class of tools and has been considered in various evaluation studies of ABE over 802.11a/b/g WLANs (*e.g.*, [39, 41]). We include pathChirp, which also belongs to the PRM category, as it has been shown to yield good results in some ABE evaluation studies in wireless network settings (*e.g.*, [107, 42]). Moreover, DietTopp and pathChirp, while from the same PRM class, differ in their probing traffic pattern, thereby allowing us to understand the relative effectiveness of different patterns. In addition, all three of them are publicly available. We briefly describe these tools in the following.

WBest. As it is explained before, WBest is a PGM-based tool and it needs to know the link effective capacity to send the probing traffic. Given that, it consists of two steps. In the first step, a number of probe packet pairs are sent to estimate the capacity. In the second step, WBest transmits a probe packet train at the estimated effective capacity rate to estimate the available bandwidth. Finally, the packet loss rate experi-

enced by the probe train is used to correct the estimated available bandwidth [40].

DietTopp. As a typical PRM based tool, this operates over multiple rounds, transmitting a probe packet train with an increased rate in each successive round. The highest sending rate with a matching receiving rate is reported as the available bandwidth estimate. DietTopp is an implementation of the measurement approach TOPP, introduced in [93] with some modification to make it less intrusive [39].

pathChirp. This employs a different probing traffic pattern called a chirp, consisting of multiple exponentially spaced probe packets of the same size, to improve efficiency and accuracy over the earlier TOPP and Pathload tools. A wide range of rates can be probed within a single chirp, thereby improving efficiency. It performs statistical analysis at the receiver to estimate the available bandwidth from multiple chirps [46].

2.6.4 Experimental Evaluation of ABE Tools in WiFi Networks

With a plethora of ABE tools, several performance evaluation studies [42, 108, 44] on these tools were conducted in the past. Lakshminarayanan *et al.* [38] are among the first who highlighted the unique challenges posed by 802.11 wireless networks' capacity and available bandwidth estimation due to use of multiple physical bit-rates, shared access and contention.

More recently, a detailed analysis of the impact of multiple access contention related delays with 802.11 CSMA/CA on active bandwidth measurements was presented in [43]. The authors of the paper introduce three reasons that make available bandwidth estimation much harder in the WLAN context: time varying capacity and MAC layer features such as retransmission and RTS/CTS. Koutsonikolas *et al.* [44] study the performance of WBest under WLAN and 1×EVDO networks.

However, none of the above works evaluate the tools against new features introduced in 802.11n networks. In Chapter 6, we characterise the behaviour of PRM and PGM tools for new enhancements in the MAC/PHY layer of the emerging WiFi networks based on the 802.11n standard.

Chapter 3

Urban WiFi Networks Characterisation Via Mobile Crowdsensing

3.1 Introduction

Smartphones are now very common handheld devices with several built-in sensors which can be seen as proper mobile sensing platforms. *Mobile crowdsensing* is defined in [70] as a technique that "individuals with sensing and computing devices collectively share data and extract information to measure and map phenomena of common interest". Similar to other proposed mobile crowdsensing applications (*e.g.*, [71, 72, 73, 74]), we consider the application of the mobile crowdsensing paradigm to wireless network monitoring using the built-in WiFi interface as the WiFi sensor. We developed a system that exploits the WiFi interface on smartphones as a sensor for low-cost and automated monitoring of WiFi networks in indoor environments like enterprises and public buildings (*e.g.*, shopping malls) and in urban areas at a city level. Specifically, we report results from a mobile crowdsensing based WiFi measurement study conducted in Edinburgh, leveraging participants with smartphones travelling on public transport buses.

Results from analysing the received data from mobile crowd sensing show that:

- WiFi spectrum usage is quite unevenly distributed across 2.4GHz and 5GHz unlicensed bands as well as among various channels within the 2.4GHz.

- Many WiFi APs contend on the same channel with around 10 other APs (and their clients) in the nearby vicinity, thereby *potentially* experiencing severe interference. This is a result of the common practice of uncoordinated and non-

adaptive channel assignment to home WiFi routers which are often left to use preset factory configuration settings for channel etc.

- We also look into the distribution of open APs, which could be leveraged for vehicular WiFi access [27]. We observe that the availability of open APs along contiguous road segments is limited to few parts near the city centre.

- We find that observations about WiFi deployments in public areas of several different indoor environments match those of WiFi deployment characteristics at city-scale.

- We validate our measurement approach by comparing it against a carefully performed wardriving study and obtain similar qualitative results.

- Our results from urban WiFi characterisation based on mobile crowdsensing are in agreement with other previous studies following different measurement approaches.

- We outline a cloud based WiFi spectrum management service for WiFi APs in urban areas (*e.g.*, home wireless routers) that can make use of results from mobile crowdsensing based urban WiFi monitoring for better interference management.

Compared to the fixed infrastructure approach (*e.g.*, Argos [22]), which relies on static deployment of WiFi monitoring sniffers, and the common practice of using wardriving [23], our mobile crowdsensing approach offers the promise of fine-grained and **continual WiFi monitoring** on a city-scale at **low cost** with comparable results to other approaches considering the limitations of the WiFi scanning approach.

The necessary background and more information about subjects discussed in this chapter can be found in Chapter 2. Specifically, §2.4 highlights the potential benefits of the mobile crowdsensing approach for urban WiFi monitoring. The rest of this chapter is structured as follows. The details of our measurement methodology are provided in §3.2. In §3.3, first crowdsensing results are verified against wardriving in §3.3.1. Then, in the rest of this section, various aspects of our measurement study, detailing mobile crowdsensing based WiFi characterisation results obtained for the city of Edinburgh, are described. §3.4 compares our results with those from related studies and discusses application scenarios of mobile crowdsensing based urban WiFi monitoring and the associated issue of incentives for user participation. A summary of this chapter can be found in §3.5.

3.2 Methodology

Our mobile crowdsensing based urban WiFi characterisation study is done using Android phones, specifically Samsung Galaxy S3 [25] phones which feature an 802.11a/b/g/n radio that can operate in both 2.4GHz and 5GHz unlicensed bands. We rely solely on passive WiFi scanning based measurements, listening to AP beacons. The information available at the user level with the Android API for passive scans is limited to: SSID, BSSID, channel, RSSI and the security scheme in use. Potentially, smartphones can be exploited as a sniffer which provides more information such as detecting clients and monitoring the volume of traffic. Practically, though, it must put the interface in promiscuous mode which is not possible with typical user access privilege.

For the measurements, we use the freely available RF Signal Tracker application for Android phones [109], which keeps passively scanning for WiFi APs in the background every three seconds or on passing 5 meters; it locally stores the result of each scan tagged with GPS location and timestamp on the phone in a text file. As this application does not log location errors and is not open source, we have a developed an auxiliary application that runs alongside and records location errors. Measurement data from phones is subsequently transferred to a back-end server where custom python scripts are used to import the data into a database, which then is used for further querying, analysis and mapping of data.

As mentioned at the outset, our urban WiFi characterisation focuses on the city of Edinburgh, which is a typical European city [110] — smaller in size compare to the USA cities and densely populated, especially in the centre.

Measurement Scenario: For proof-of-concept and wider spatial coverage with fewer participants in a short measurement period, we focus on a measurement scenario where participants are travelling on public transport vehicles. Specifically, our measurement results are obtained from phones carried by participants during the times they travel at low to moderate speeds on buses in the city operated by a local bus company called Lothian Buses [111]. In this sense, it follows a participatory sensing approach along the lines of earlier urban air/noise pollution monitoring studies [72, 74].

Measurements reported in this chapter correspond to travelling on 31 buses over a 15 hour period in total. Note that in principle crowdsourcing based measurement can be done in a fully opportunistic manner, covering all modes of movement including walking, standing, etc. The limits we place are for the above mentioned reasons. Also

Figure 3.1: Mobile crowdsensing based WiFi AP scanning measurements shown as a heatmap.

Table 3.1: Location error statistics for the collected measurement dataset.

	Min	Median	Mean	Max
Location Error (m)	4	8	9.6	1095

note that there is an assumption underlying our study that APs visible from next-door neighbours can also be seen from the street and vice versa.

Measurement Statistics: Figure 3.1 shows the total set of measurements as a heatmap. Red areas in the map indicate places where there is a high density of APs as well as those places with multiple measurements due to overlapping road segments between different bus routes. Table 3.1 lists the location error statistics across all measurements in our dataset. We observe that while the maximum error can be over 1km reflecting locations that do not get a GPS fix, the error is under 50m in 95% of the cases. To obtain reliable spatial distribution of APs on the map, we filtered out the 5% of measurements with location errors greater than 50m. Table 3.2 presents a summary of the resultant dataset. From closer inspection, we observe that the majority of the APs correspond to

Table 3.2: Filtered measurement dataset summary.

Total number of measurements (scans)	147488
Distinct measurement locations	11225
Distinct APs detected	13800
Distinct open access APs detected	2977

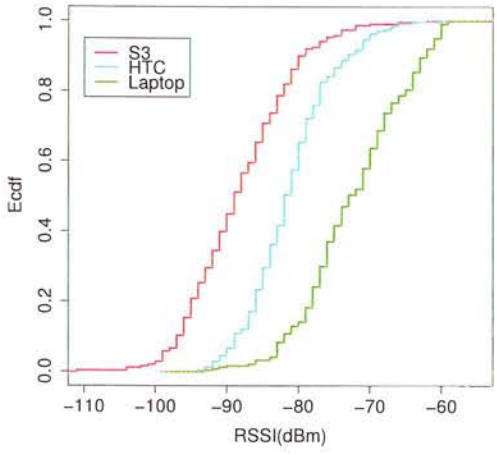


Figure 3.2: Empirical CDF of maximum RSSI for common WiFi networks seen across different measurement devices.

home WiFi networks interspersed with the rest (*e.g.*, WiFi hotspots).

3.3 Results

In this section, we analyse captured data through mobile crowdsensing to verify the proposed approach against wardriving and show information that can be harvested from sensing data, specifically to look into mutual interference characterisation.

3.3.1 Comparison Crowdsensing with Wardriving and Device Effect

To validate the mobile crowdsensing approach, we compared it against a laptop based wardriving study. Specifically the validation experiment was carried out over a Edinburgh University shuttle bus that connects two university campuses (one in the city center and the other in south Edinburgh), 2.7Km apart. For the wardriving part of the

experiment, we used a customised Lenovo T420 laptop with GPS and running only inSSIDer WiFi scanning software [64]. Two different smartphones, Samsung Galaxy S3 and HTC (Google Nexus One), both running the RF Signal Tracker application in the background were used for mobile crowdsensing. Note that all other measurement results reported in this chapter were obtained with Samsung Galaxy S3 phones while Google Nexus One phone is only used in this experiment to better understand the device effect.

During the journey, 429 APs were detected by inSSIDer with the laptop while Galaxy S3 and Nexus One could detect 384 and 404 APs respectively. This shows that the commodity smartphone based mobile crowdsensing approach can detect nearly all ($> 90\%$) APs that can be seen by the wardriving laptop. This is remarkable considering that laptops are equipped with better antennas and radios with higher receive sensitivities, a fact confirmed by higher RSSI values obtained with a laptop in the experiment (Figure 3.2). RSSI values for the two phones indicate device diversity (in terms of radio, antenna and platform design) and partly explain differences in the number of networks detected between them (this problem is studied previously in the context of WiFi fingerprinting [112]). Note that some of the differences in scanning results between the three cases stem from differences in channel hopping sequence and duration between different devices and software, which are outside our control in all three cases compared (the hopping sequence and duration for a complete channel list scanning depends to the number of channels supported by the device. For example, the Samsung S3 phone, the laptop and the HTC one phone support different number of channels, where S3 and the laptop also support channels in the 5GHz band, the HTC one phone only supports channels in the 2.4GHz band. Note that the results reported in §3.3.1 are only representative for APs detected in the 2.4GHz, to ensure consistency between all devices under testing).

Overall, the results from this experiment demonstrate that mobile crowdsensing with commodity smartphones can yield similar results to those obtained via carefully conducted wardriving campaigns.

Given acceptable performance from mobile crowdsensing approach, in the following sections we present results of analysing and visualising data received from mobile phones. Specifically, we investigate how ISM spectrum freely available for WiFi usage is shared between un-planned WiFi networks across the city and show their spatial distribution through illustration on the map.

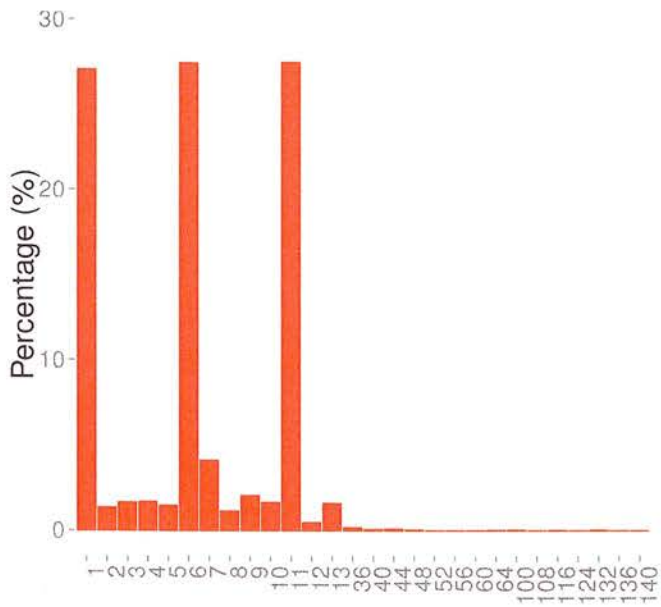


Figure 3.3: Relative usage of different channels across the 2.4GHz and 5GHz bands by the detected APs.

3.3.2 Spectrum Usage

We begin by looking at the channel usage of WiFi APs in our dataset. Figure 3.3 shows the relative usage of different channels across the 2.4GHz and 5GHz bands. Clearly, the channel usage is quite uneven, dominated by channels 1, 6 and 11 from the 2.4GHz band. We attribute this primarily to users leaving their APs to use factory settings, which commonly focus on channels 1, 6 and 11 given that they are non-overlapping. Among the rest of the channels, channel 7 is the next most common channel, which we find is due to the fact that WiFi APs corresponding to one of the ISPs (identified based on their SSID) are always set to use channel 7. The very little perceived use of 5GHz channels may be partly due to the relatively poor propagation characteristics at 5GHz and our measurement from outdoors while APs are almost always located indoors. Nevertheless, we do not expect our conclusion on the unevenness of channel usage to change qualitatively given the results discussed later in this section on the nature of WiFi deployments seen in different indoor environments.

We explore this observed non-uniform channel use further in the next subsection, particularly looking at spatial variation in spectrum usage and its implication for potential interference levels.

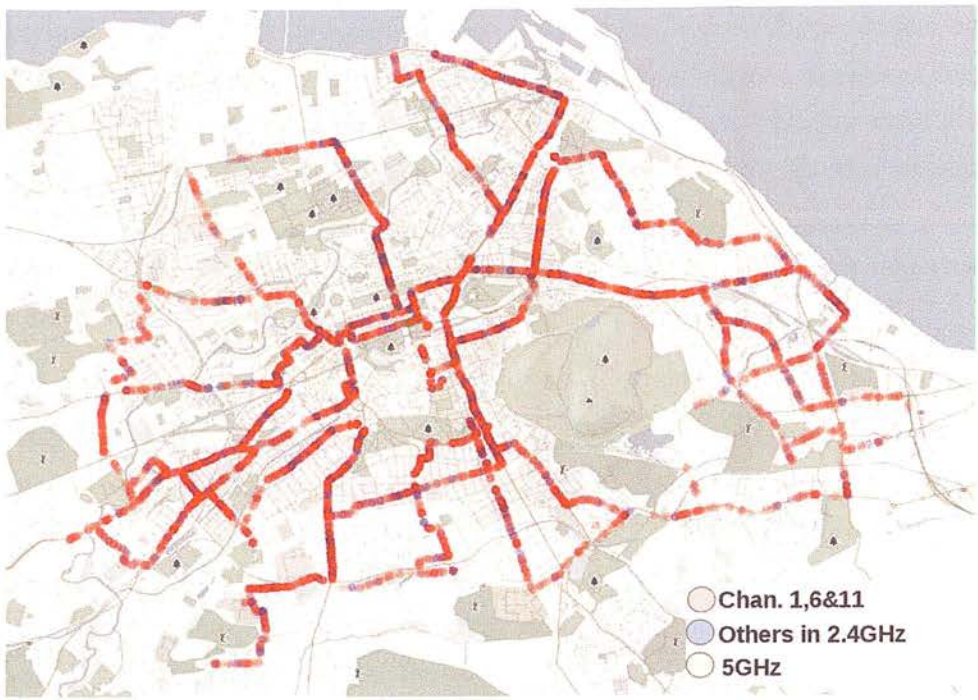
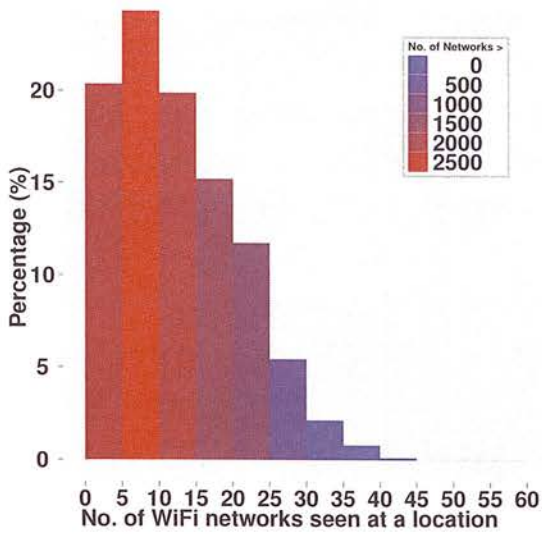


Figure 3.4: Map of distinct APs detected.

3.3.3 Spatial Distribution of Spectrum Usage

Figure 3.4 shows the map of detected APs, coloured differently depending on the set of channels used. Besides confirming the channel usage pattern from Figure 3.3, the red patches on the map highlight the closeness between APs using one of three popular channels (1, 6 and 11), and thereby the potential for high interference. Figure 3.5 provides a quantitative equivalent of the map in Figure 3.4 and shows that more than half of the locations "see" more than 10 APs (statistics in the table from Figure 3.5 confirm the same).

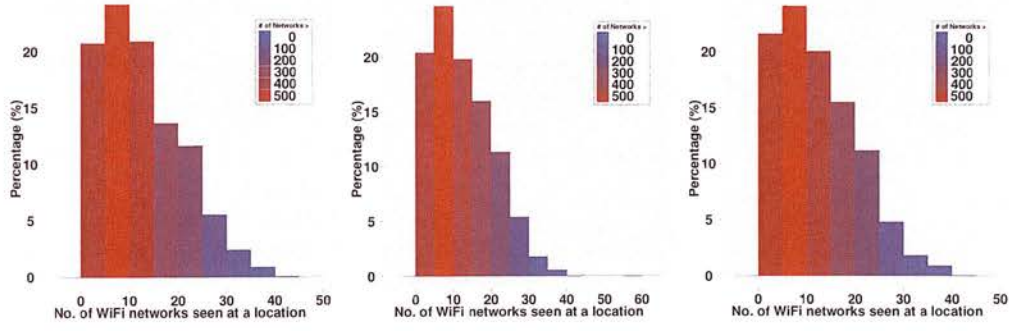
Since Figure 3.5 corresponds to the spatial distribution of AP density *over all channels*, it does not directly represent levels of interference in any one channel. This information is shown in Figure 3.6 for the three mostly commonly used channels (1, 6 and 11). The striking aspect is that the spatial distributions of AP density for each of these channels are similar to the aggregate distribution spanning all channels shown in Figure 3.5. Hence we can actually infer that in over half of the locations we are likely to find more than 10 APs on any of the three heavily used channels. The same result is illustrated in Figure 3.7 on a map. It shows the locations with 10 or more APs configured to use the most common channel at that location with the size of each circle representing the number of APs at a location — larger the size of the circle,



	# APs
Minimum	1
1st Quartile	6
Median	12
Mean	13.4
3rd Quartile	19
Maximum	59

Figure 3.5: AP density spatial distribution (left); AP density statistics across all measurement locations (right).

more the number of APs that could *potentially interfere* with each other (or their associated clients). It is worth to recall the previous studies (*e.g.*,[54, 113]) that shows APs operating in a same channel and in the transmission range of each other can interfere and severely affect the performance of QoS sensitive applications like VoIP. Note that the amount of concurrent traffic generated by each AP cannot be monitored by WiFi scanning and as mentioned before, it is the limitation of this measurement technique.



(a) Channel 1 (b) Channel 6 (c) Channel 11

Figure 3.6: Per channel spatial distribution of AP densities.

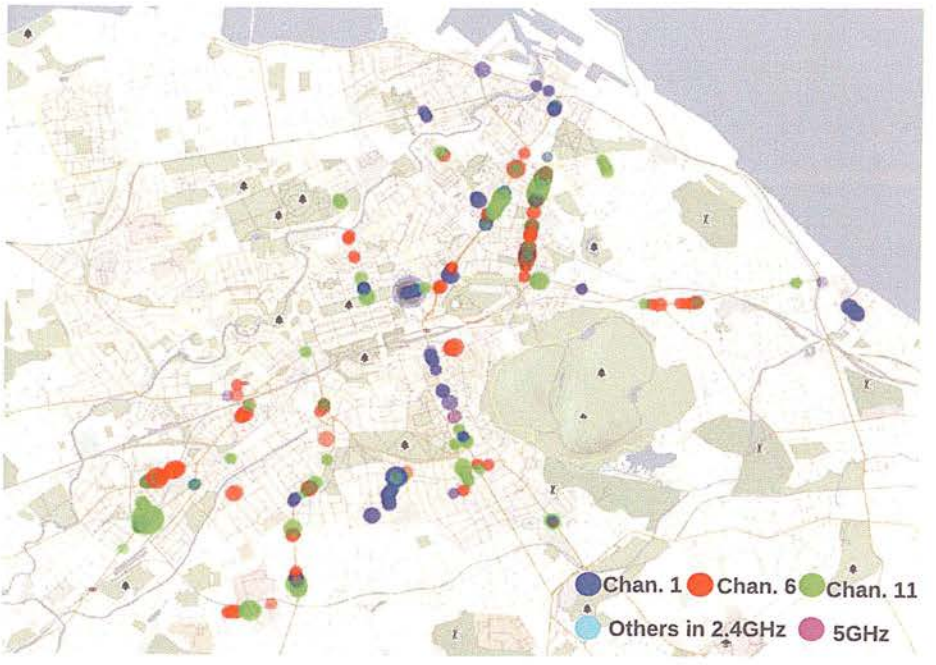


Figure 3.7: Map illustrating likely high interference locations (with more than 10 mutually interfering APs).

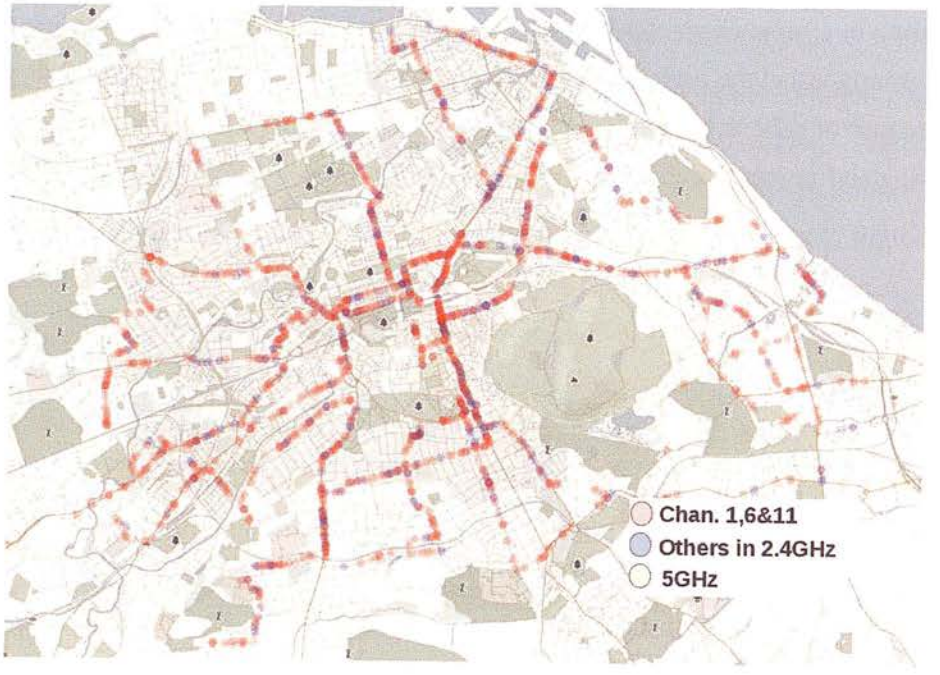


Figure 3.8: Map of open APs detected.

3.3.4 Open Access Points

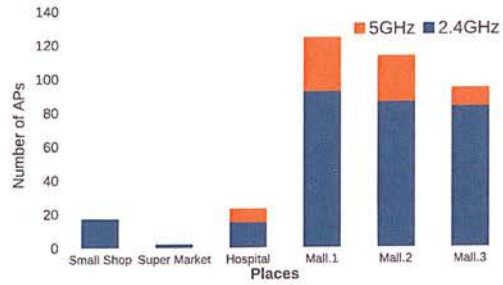
Here we look into the question of "open" APs which could be exploited for public and vehicular wireless Internet access in cities. Open APs are those APs that support open-system authentication *i.e.*, stations associate to the AP without being authenticated, however, authentication might be carried out after being associated to the AP, *e.g.* web authentication in Fon WiFi [16].

In our measurement dataset, we find that open APs constitute around 20% of the total number of APs detected (2977 vs. 13800). And a large fraction of these open APs (nearly 76%) are served by a single ISP — British Telecom (BT), making it plausible to view them all to be part of a single administrative domain from a vehicular client perspective for seamless roaming. This argument is made stronger by the fact that BT in the UK has a partnership with the Fon WiFi community network [16], making every BT broadband customer automatically a member of the Fon network. However, the spatial distribution of open APs along roads (Fig. 3.8) suggests that the presence of a contiguous set of APs with overlapping coverage areas is limited to few areas in the very centre of the city, limiting the possibility of seamless vehicular WiFi Internet connectivity via open APs.

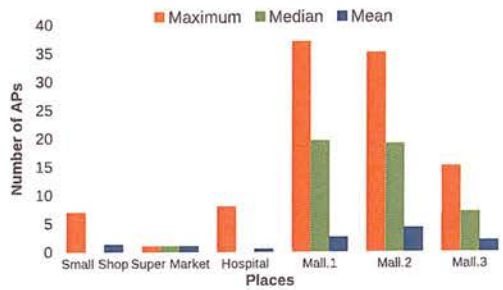
3.3.5 Comparison with Indoor Environments

We study the characteristics of public WiFi deployments in indoor environments as a way to increase the confidence in our findings from outdoor measurements concerning the nature of urban WiFi networks. For this purpose, we developed a custom mobile application called IndoorScanner based on Funf [114]. The need for a different measurement app for indoors is motivated by the fact that GPS does not reliably work indoors and given that the RF Signal Tracker app used for our outdoor measurements relies on GPS for locating measurements. In contrast, IndoorScanner requires the user to select the measurement location on a digital map of the indoor environment (*e.g.*, floor map, building layout) in a manner similar to traditional site survey procedures for WiFi fingerprinting based localisation systems. Note that these indoor measurements were one off and gathered by a single user, hence we did not need to employ Pazl [81] for this purpose.

We consider several different indoor environments located in different parts of the city for this study. These include: three different shopping centers, a large hospital, a supermarket, and a small shop. We carefully measure in public places inside these



(a) Spectrum usage



(b) Max AP density per channel

Figure 3.9: WiFi scanning measurement results from different indoor environments.

environments looking for the presence of WiFi networks. As shown in Figure 3.9(a), WiFi use in indoor environments happens largely in the 2.4GHz band just as is seen from outdoor measurements (cf. Figure 3.3). Figure 3.9(b) shows that the maximum number of APs at a location using the same channel can be as high as 37, which has been also observed from outdoor measurements (see Figure 3.6).

3.4 Putting Our Findings Into Perspective

In this section, we compare our findings with other related studies on WiFi characterisation and end-user performance assessment in the urban context.

Akella *et al.* [115] analyse several wardriving datasets and observe that up to 85 APs could be within close proximity of each other for an assumed interference range of 50m. They also find that more than 40% of APs are configured on channel 6 in one of the datasets. Our results are qualitatively similar but obtained using a different, mobile crowdsensing, approach.

Two recent studies reported in [5] and [4], both commissioned by the UK communications regulator Ofcom, are closely related to our work in terms of the underlying

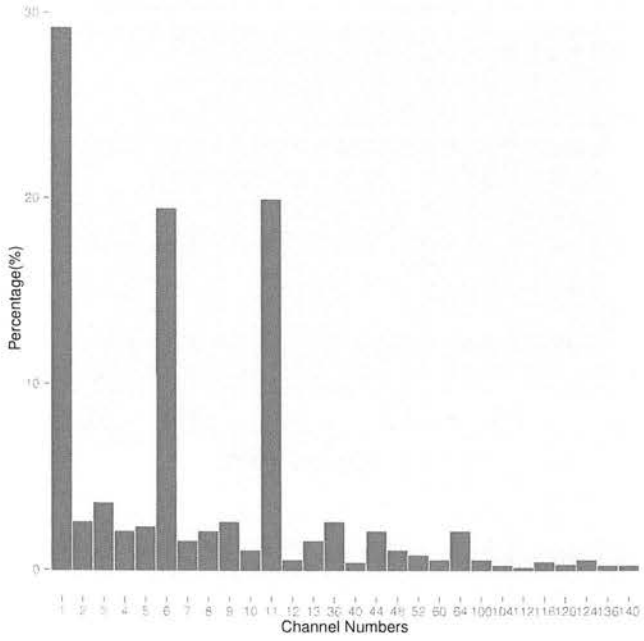


Figure 3.10: Number of APs detected in different channels across 2.4GHz and 5GHz bands with walk around survey in *central London* [4].

goals to characterise WiFi usage in urban areas across the unlicensed 2.4GHz and 5GHz bands and in different environments. Recall from our discussion in §2.4 that these studies use approaches that are different from the mobile crowdsensing approach we take ([5] relies on a fixed measurement infrastructure, whereas [4] is wardriving based). Nevertheless, they report observations similar to our findings described in the previous section. We elaborate on some of these below for concreteness.

In [4], WiFi channel usage measurements across 2.4GHz and 5GHz via walk around surveys in central London neighbourhoods show that the majority of APs are configured to one of the three non-overlapping channels (1, 6, 11) in the 2.4GHz band as shown in Fig. 3.10. This is similar to what we also found in Edinburgh although with a different measurement approach (cf. Fig. 3.3). An additional interesting observation made in [4] is that public WiFi hotspots are deploying their APs in 5GHz channels, which suggests the increased use of the 5GHz band in future.

[5] studies the usage in 2.4GHz and 5GHz unlicensed bands and WiFi performance in different environments (houses, apartments, cafés and shopping centres) with the help of fixed installations of monitoring equipment at various selected locations. Similar to our study, it concludes that the 2.4GHz band is more heavily occupied (10 times

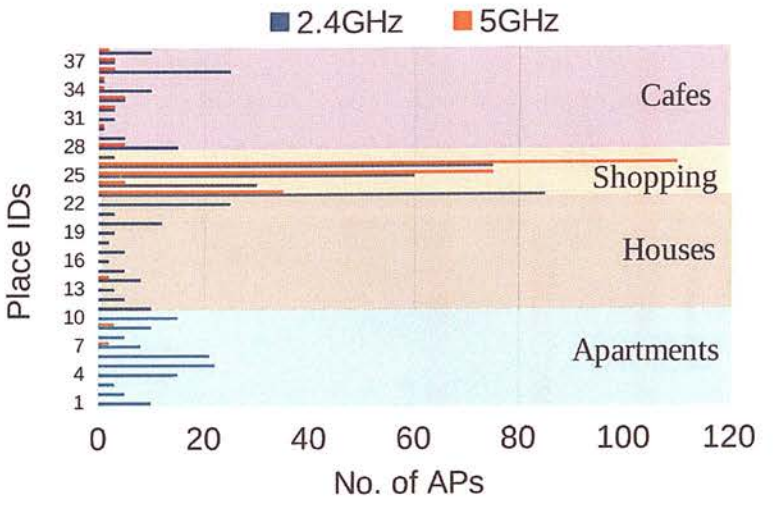


Figure 3.11: AP densities across different environments as reported in [5] via fixed monitoring kit at different locations.

or more) than the 5GHz band; it identifies this to be mostly due to WiFi transmissions in 2.4GHz and not because of other types of 2.4GHz usage such as Bluetooth, ZigBee and microwave ovens. It also has similar conclusions about rather high AP densities in some cases. Fig. 3.11 shows a sample of the results from [5] for reference.

The above discussion attests to the validity and reliability of commodity smartphone based mobile crowdsensing approach for urban WiFi characterisation and monitoring. It is even more remarkable that we are still able to obtain similar conclusions despite the inability to obtain lower level metrics such as channel utilisation and number of MAC retransmissions with the current APIs on smartphones.

There also exist several studies that examine the negative impact of unplanned and uncoordinated urban WiFi deployments on end-user performance (*e.g.*, [8, 115]), and those that investigate optimised AP configuration (channel, transmit power, etc.) and association mechanisms to mitigate such performance degradation (*e.g.*, [26]). Given our observations concerning high density of APs in some locations, the analyses on the impact of high AP densities with unplanned WiFi deployments on end-user performance are particularly relevant. Recall from Chapter 2, the authors in [8] experimentally investigate the effect of AP density (equivalently, inter-cell interference) and client density on performance of different applications such as web and multimedia using the ORBIT testbed [9]. Their results show that increasing the number of clients to 125+ in a single AP WiFi deployment scenario does not degrade the collision rate

and throughput much, which is similar to what is reported in [116]. In contrast they find that in an unplanned multi-AP WiFi deployment scenario increasing the number of APs causes a significant increase in collision rate and consequent high drop in throughput; for example, aggregate throughput drops by 50% with only four interfering APs with the same overall number of clients as in a single AP scenario. Media streaming performance is also seen to take a big hit in the presence of inter-cell interference. For the voice over IP (VoIP) application, substantial performance degradation is seen in the multi-AP scenario with just three APs — average latency increases from 54ms in the single AP scenario to 304ms in the scenario with four uncoordinated APs; jitter also increases four-fold with the multi-AP scenario.

3.4.1 Applications for Mobile Crowdsensing Based Urban WiFi Monitoring

The findings from our measurement study and the foregoing discussion suggests that unplanned and uncoordinated home or hotspot WiFi networks in urban areas can potentially suffer from severe interference related performance degradation. This can be seen as real world evidence showing that vast research on self-organisation mechanisms for channel and transmit power allocation in unplanned WiFi deployments (*e.g.*, [26]) has not actually materialised. We observe that the impediment for large-scale deployment of intelligent self-organisation mechanisms in practice may not be technical but rather the lack of market incentives for their application.With this in mind, we outline an alternative approach that may find greater real world acceptance.

The idea is for a mobile crowdsensing based urban WiFi monitoring system to continually feed spectrum usage measurements to a cloud based back-end, which takes the global awareness of spectrum usage and interference conditions to determine the best channel for each participating WiFi AP (home WiFi router). Such a spectrum management service could be subscription based and tied to the user's broadband service plan — the user's home WiFi AP can be reconfigured on the fly via the ISP, informed by the cloud based spectrum management service (see Figure 3.12). Such managed and coordinated spectrum management approaches are emerging in other related domains such as efficient sharing of TV white space spectrum among secondary users (see [117], for example).

Another application scenario for mobile crowdsensing based urban WiFi monitoring is targeted toward outdoor small cell public WiFi based hotspots run by several

different operators. The deployment of such hotspots is experiencing a high growth and is seen to complement LTE small cells in an overall solution to aid in better managing the steeply rising mobile data traffic[1]. The emerging passpoint technology [118] to enable seamless roaming between public WiFi hotspots run by different operators will play a role in their widespread deployment and use, and in turn determine the need for coordinated interference management.

Concerning incentives for user participation in mobile crowdsensing based urban WiFi monitoring, real world evidence suggests that smartphone users have sufficient incentives to participate in crowdsourced mobile network measurement campaigns. For instance, in a 3G crowdsourcing measurement study [119] conducted by BBC in partnership with measurement firm Epitiro, nearly 45,000 volunteers installed the measurement app to participate within one month of announcing of the study. As another example, OpenSignal [78], another firm, with an app for crowdsourced mobile measurement has over 3 million users in over 200 countries collectively reporting over 4 billion measurement samples to date. If such voluntary participation exists, the offer of better connectivity or cheaper service while on the move may provide an incentive for mobile (smartphone/tablet) users to participate. We note that devising suitable incentives for mobile crowdsensing is a topic in itself and is currently receiving a lot of attention in the research community.

3.5 Summary

We proposed a mobile crowdsensing approach for the first time for large scale WiFi deployments. We have shown the value of this approach for urban WiFi characterisation and monitoring through a measurement study in the city of Edinburgh. Our results indicate that the uncoordinated and inefficient spectrum use is the source of potentially severe interference problems that might be seen in practice at locations with high AP densities. We have also found similarity between our outdoor city-scale WiFi measurement results and characteristics of WiFi deployments in several different indoor environments. We have validated our approach against a carefully conducted wardriving journey. Our results and findings are also in agreement with other previous urban WiFi characterisation studies based on other measurement approaches. Finally we have outlined a cloud based spectrum management service that could leverage results

[1] Europe loves Wi-Fi: new study recommends more spectrum should be made available (http://europa.eu/rapid/press-release_IP-13-759_en.htm)

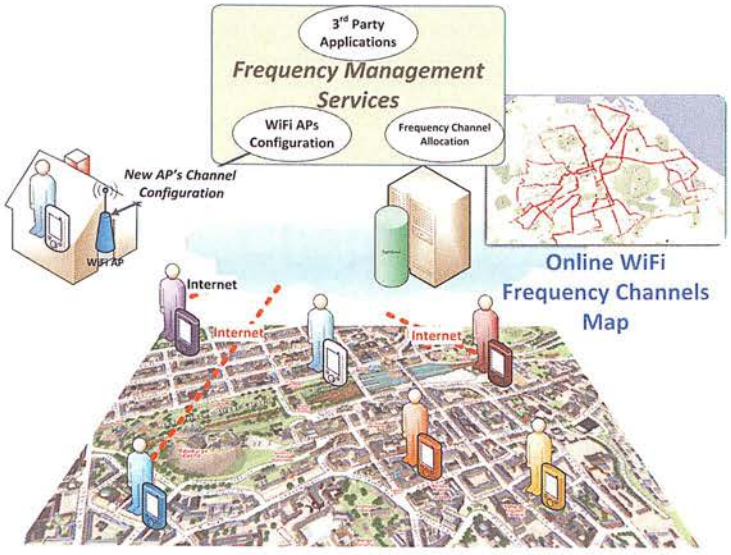


Figure 3.12: Cloud-based crowdsensing WiFi spectrum management system.

from mobile crowdsensing based urban WiFi monitoring for more effective interference management in urban WiFi networks.

Chapter 4

Coexistence Interference Characterisation and Mitigation

4.1 Introduction

In a multi-radio platform (*e.g.*, multi-radio mesh router), individual APs/routers are equipped with multiple radio interfaces, each configured to different channels. The advantages of such an architecture for achieving high end-to-end network performance, keeping (co-channel) interference low and better utilisation of available spectrum are well established, both theoretically and through simulation-based evaluations of channel allocation protocols (*e.g.*, [120]). Realising the benefits of multi-radio architectures in practice poses challenges that are often abstracted out in simulation-based evaluations. These concern interference resulting from collocation and simultaneous operation of multiple radio interfaces within a platform. Following [15], we will refer to this interference as *multi-radio coexistence interference*.

As introduced in §2.2.2, this interference is a composition of three phenomena: *receiver blocking*, *transmitter noise* and *intermodulation*. Receiver blocking is a result of the limited dynamic range of the power amplifier and A/D converter in the receiver; it arises in situations where a transmitting (interferer) antenna is in close proximity to a collocated antenna. If the total input power at the receiver is more than the blocking limit (*e.g.*, -30dBm at 2.4GHz for WiFi) the received signal strength degrades. Transmitter noise or adjacent channel interference (ACI) refers to the out-of-band emission seen by receivers in close proximity to a transmitter (*e.g.*, due to imperfect filtering at the transmitter antenna). Intermodulation is a result of non-linearity of radio components such as the amplifier. It surfaces when intermodulation bandwidth due to

a pair of concurrent and nearby transmissions overlaps with receiver channel bandwidth. We experimentally show that performance degradation due to coexistence interference on multi-radio 802.11 platforms can be significant if adequate care is not taken to mitigate it. Physical separation of antennas on 802.11 multi-radio platforms and separation between channels used by different collocated radio interfaces are common ways to avoid performance degradation due to multi-radio coexistence interference [13, 28, 29, 14, 30, 31]. The required amount of separation between antennas' positions and interfaces' channels varies depending on factors such as transmission power and bit-rate. Ideally, we would want these separation amounts to be as small as possible while allowing us to transmit at maximum bit-rates allowed by link/channel quality and at maximum transmit power up to the regulatory limit. Small antenna separation will lead to a compact platform that is easier and cheaper to deploy, an especially important consideration for indoor WLAN/mesh scenarios. Smaller channel separation can potentially allow better utilisation of the available spectrum. Our experiments show a somewhat surprising result: multi-radio coexistence interference (due to receiver blocking) can be so severe that even collocated radio interfaces operating on channels from *different bands* (*e.g.*, 2.4GHz and 5GHz) interfere with each other when their antennas are in close proximity. This observation holds true across different multi-radio platforms and 802.11 interface cards.

To relieve the need for increased antenna/channel separation without limiting transmit power/bit-rate, we consider *antenna polarisation*, for the first time, as an extra knob to introduce an additional coupling loss up to 20dB between collocated antennas. Coupling loss refers to the amount of drop in strength of interference to alleviate multi-radio coexistence interference. If it is greater than the *minimal coupling loss* then coexistence interference can be eliminated [15]. Polarisation is the direction of the electric field of a radio wave relative to the ground. Linearly polarised antennas are commonly used in 802.11 networks and they are typically polarised either vertically or horizontally — the electric field is perpendicular to the ground for vertically polarised antennas, whereas it is parallel to the ground with horizontally polarised antennas. Antenna orientation and polarisation are closely related in the sense that changing the orientation of the antenna changes its polarisation.

- We experimentally show that having differently polarised antennas for different 802.11 interfaces on a multi-radio node reduces required antenna separation to as low as 3cm in multi-band configurations.

- When widely separated channels within a single band are used for different collocated radio interfaces, up to four times higher bit-rates are made possible by differently polarised antennas compared to using identically polarised antennas for the same antenna separation and transmit power.

- On the other hand, when nearby channels are used between different interfaces on a multi-radio 802.11 platform, the required amount of channel separation to avoid performance degradation is reduced when interfaces use differently polarised antennas.

- Crucially, the above benefits do not come at the cost of network connectivity. Using measurements from an actual indoor multi-radio mesh network testbed, we show that the use of differently polarised antennas has a small effect on the mesh network topology and link qualities. Essentially, we exploit the fact that after a few reflections the polarisation of the signal at the transmit antenna does not have a bearing on the polarisation of the signal at the receiver side [121]. This is particularly true in non-line-of-sight environments as also experimentally demonstrated in previous work [122].

It is important to note that omnidirectional antennas that allow changing polarisation are only slightly more expensive compared to those that do not offer such flexibility. Compared to previous work (discussed in §2.2.2), the main contributions of this study lie in:

- Highlighting the severity of multi-radio coexistence interference due to receiver blocking even when using different frequency bands.

- Demonstrating the use of antenna polarisation as a means to alleviate multi-radio coexistence interference with few negative side effects.

- Characterising adjacent channel interference using direct observation of MAC behaviour and validation of observations from prior work based on indirect packet delivery ratio measurements (*e.g.*, [14]).

The necessary background and more information about subjects discussed in this chapter can be found in Chapter 2. Specifically, in §2.1.1 multi-radio platforms are introduced. Coexistence interference and previous related works are discussed in §2.2.2. The rest of this chapter is structured as follows. In §4.2, the experimental methodology

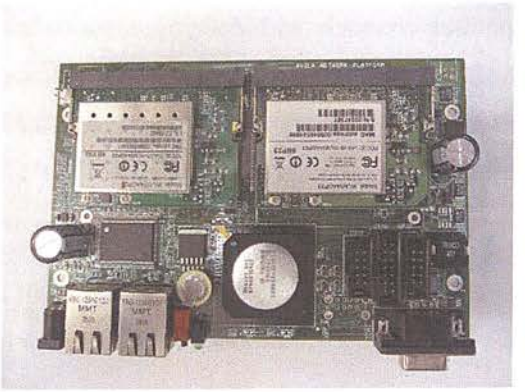


Figure 4.1: Gateworks Avila multi-radio platform equipped with two Compex WLMAG54-23dBm mini PCI cards.

including details on testbed setups and equipments to build multi-radio platforms is explained. §4.3 studies coexistence interference for two cases: *multi band case*, when interfaces are operating in two different bands (§4.3.1) and *single band case*, when interfaces are configured in different non-overlapping channels from the same band (§4.3.2). A direct view of the ACI effect from the MAC/PHY layer is shown in §4.3.3. In §4.4, the effect of using a differently polarised antenna on the network connectivity is experimentally studied. §4.5 summarises findings and contributions of this study.

4.2 Methodology

Several factors influence the nature and extent of multi-radio coexistence interference, including: *platform, types of radio interfaces and antennas, antenna and channel separation, transmit power and bit-rate.*

When characterising coexistence interference these factors and their mutual interaction have to be considered. For our experimental study, we consider three different multi-radio platforms: Gateworks Avila [123], Gateworks Cambria [124] and Ubiquiti RouterStation [125]. As we found similar results using these different platforms, we only present results for experiments based on the Avila platform.

For the operating system on the platform, we use OpenWrt Linux. For 802.11 radio interfaces, we use two different types of miniPCI cards: Compex WLM54G-23dBm [126] and Mikrotik R52Hn [127]. As with the platform, we mainly present results for experiments using Compex cards.

Figure 4.1 shows the Avila platform with two Compex cards installed. Both Com-

Table 4.1: Experiment Hardware and Software.

Router Platforms	Gateworks Avila and Cambria, Ubiquiti RouterStation
Platform OS/Firmware	Openwrt attitude r27891
Radio Interfaces	Compex WLMAG54-23dBm and MicroTik R52Hn mini PCI cards
Device Drivers	ath5k / ath9k r2011-06-22
Antennas	Cisco dual-band dipole omni (2dBi@2.4GHz, 5dBi@5GHz) Laird dualband omni (3dBi@2.4GHz, 5dBi@5GHz)

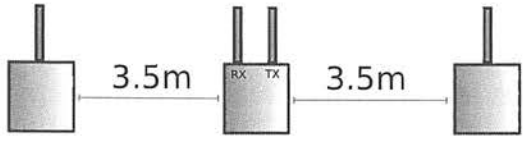


Figure 4.2: Experiment setup with a dual-radio node in the middle.

pex and Mikrotik cards used in our study are based on Atheros chipsets. We use open-source ath5k/ath9k device drivers. The Mikrotik card is actually a 2x2 MIMO supporting the 802.11n standard but we use it in legacy mode.

Previous studies have reported that Atheros-based cards have some undocumented features such as Ambient Noise Immunity (ANI) and antenna diversity [128]. We experimentally found that they are not helpful in mitigating multi-radio coexistence interference; as such they are not relevant for this study.

We also use two types of dual-band omnidirectional antennas, one from Cisco and the other from Laird Technologies, with Cisco antennas [129] being our default, as they allow changing of antenna orientation (polarisation). The above settings are summarised in Table 4.1. Other parameters such as antenna separation and transmission power are varied in our experiments.

As in previous studies on characterisation of coexistence interference on multi-radio 802.11 platforms [13, 28, 14, 30], we use the three node experimental setup as shown in Figure 4.2. The middle node in the setup is equipped with two radios and the performance of transmissions/receptions on those radios while varying other parameters (*e.g.*, transmission power, antenna separation, channel distance, etc.) is the main focus of the study. In terms of the traffic workload, we use Iperf tests and measure the UDP throughput. We also study other metrics derived using or related to throughput such as minimum antenna separation and maximum achievable bit-rate.

4.3 Multi-Radio Coexistence Interference

In this section we study coexistence interference in two setups: when interfaces are configured in multi-band and when they are configured in different channels of a band.

4.3.1 Multi-Band Case

In multi-band configuration, each node is equipped with two radios configured to use channels in two different frequency bands. Towards this end, we use the three-node setup from Figure 4.2 with the middle dual-radio node having 2 Compex cards, each configured to a channel in 2.4GHz and 5GHz unlicensed bands, respectively. The 2.4GHz interface is used to transmit Iperf UDP traffic to an end node while the 5GHz interface receives identical traffic from the other end node. Although the traffic direction is found to make a difference as previously observed in [30], results are qualitatively similar and do not affect our conclusions. Transmission power on the 2.4GHz interface is varied over a wide range from 5dBm to 23dBm while keeping the bit-rate on the interface fixed to 6Mbps. Our focus is on measuring the extent to which data can be successfully received on the 5GHz interface. We use two extreme bit-rates of 6Mbps and 54Mbps for the 5GHz link to capture the reception performance at widely different transmission rates. We define a metric called *minimum antenna separation (in cm)* that corresponds to the smallest antenna separation between transmitting and receiving interfaces in the dual-radio node that yields closest to maximum throughput for the bit-rate in question (around 5Mbps for 6Mbps bit-rate and 29Mbps for 54Mbps bit-rate). This metric directly captures the impact of multi-radio coexistence interface on the platform size. Note that, this metric is measured in a single experiment by varying the antenna separation until achieving the minimum separation without loosing the throughput performance.

Results are shown in Figure 4.3 focusing on the typical configuration where multiple collocated radios use identically (vertically) polarised antennas, we observe that the minimum antenna separation increases with the increase in transmit power of the other transmitting (2.4GHz) radio and increase in reception bit-rate of the receiving (5GHz) radio. Both these are along expected lines — higher transmit power causes more interference, whereas reception at higher bit-rate requires higher SINR (or alternatively, lower interference). However, the fact that this happens even with collocated interfaces using widely differing frequency bands is quite remarkable. We attribute this behaviour to the receiver blocking effect. We obtained greater confidence in this

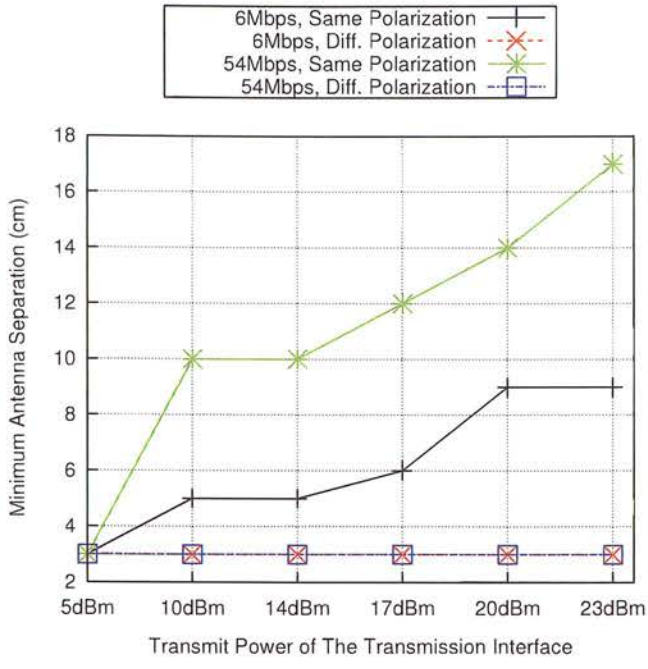


Figure 4.3: Minimum antenna separation required to avoid throughput degradation on Avila-based dual-radio 802.11 platform at different bit-rates, antenna polarisations and collocated interferer transmit power levels.

conclusion when we found that the MAC on the receiving (5GHz) interface never got into BUSY state [130] when the collocated 2.4GHz interface was transmitting.

In order to reduce the undesirable interference from a collocated transmitting radio, we experiment with having the two collocated antennas use different polarisations (one using vertical and the other horizontal). Results for this configuration are also shown in Figure 4.3 and are significantly different from the earlier results. Minimum antenna separation now always remains at the minimum *3cm* and is unaffected by the increase in transmit power of the collocated interfering interface regardless of the bit-rate used by the receiving interface. This shows that the additional coupling loss introduced by using differently polarised collocated antennas is sufficient to have the two links function concurrently without hurting each other and, more importantly, without needing increased antenna separation (platform size).

In order to confirm that the above results are not peculiar to the specific type of hardware used, we experiment with a different interface card (MikroTik), using different cards in combination (MikroTik and Compex) and with other platforms (Cambria and Ubiquiti RouterStation). Corresponding results when using different cards are

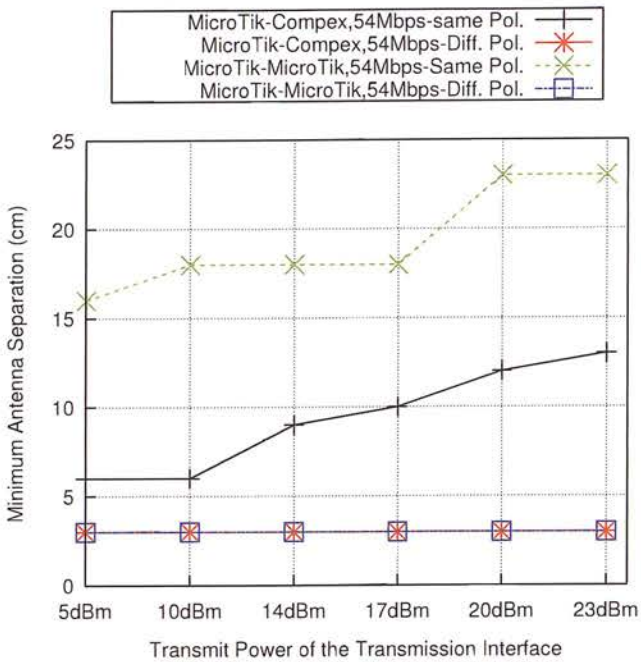


Figure 4.4: Impact of using different radio interface cards in combination on minimum antenna separation required to avoid throughput degradation on Avila-based dual-radio 802.11 platform for the 54Mbps bitrate, antenna polarisations and collocated interferer transmit power levels.

shown in Figure 4.4. Although the absolute numbers increase a bit for the case with identically polarised antennas, results are qualitatively similar to Figure 4.3 and previous conclusions still hold. Note that results for 6Mbps transmission for different WiFi interfaces are exactly the same as for the identical WiFi interfaces; so, their results are not illustrated (the minimum antenna separation of 3cm is observed for different transmission power levels and different WiFi interface combinations). We did not observe any noticeable difference in the results by changing the platform (not shown).

We now look at the impact of multi-radio coexistence interference on received UDP throughput in different configurations. For this, we use the same setup with Compex cards as before but set the transmit power of the transmitting (2.4GHz) interface to 17dBm, the default for Compex cards, and fix the antenna separation to 6cm, which from Figure 4.3 is the minimum required to avoid degradation at 6Mbps bit-rate when the interfering radio transmit power is 17dBm and identically polarised antennas are used. Unlike before where we only considered the extreme rates, we now measure the received throughput at *all* bit-rates as shown in Figure 4.5. These results confirm

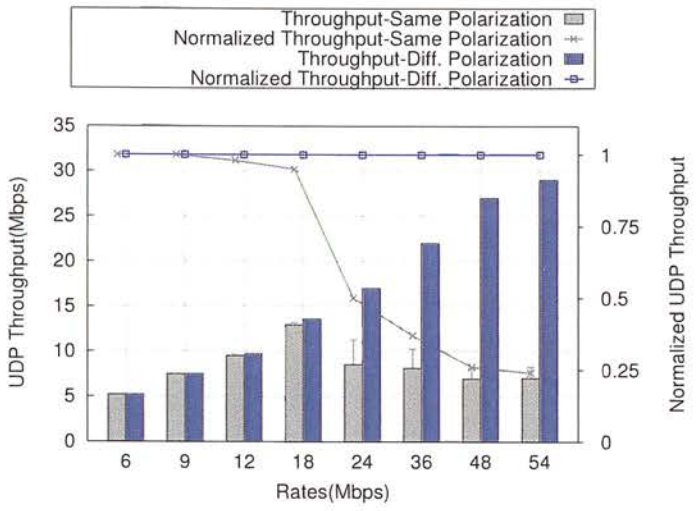


Figure 4.5: Received UDP throughput as a function of different bit-rates when antenna separation is fixed at 6cm and power level of the interfering transmit interface on 2.4GHz is set to 17dBm.

that using differently polarised antennas always yields maximum throughput, whereas maximum throughput is achieved only up to about 18Mbps bit-rate with the typical configuration using identically polarised antennas and throughput degrades beyond that point, only providing less than a quarter of the maximum throughput at the 54Mbps rate.

The overall conclusion from our study of dual-radio platform in multi-band configuration is that *the use of differently polarised antennas is a very effective remedy to counter multi-radio coexistence interference and achieve high performance without increasing platform size.*

4.3.2 Single-Band Case

So far, we have studied multi-radio coexistence interference when multiple radios operate in different bands. We now consider the case where different radios in a multi-radio platform are configured to use channels from within a single frequency band, focusing on the 5GHz unlicensed band that has a number of channels making it suitable for 802.11-based multi-radio networks.

We first consider the situation where two radios on a dual-radio platform are configured to use *far apart channels within the same band.* Specifically, we assign channels 36 (5.18GHz) and 157 (5.785GHz) to the two radio interfaces on the middle node in

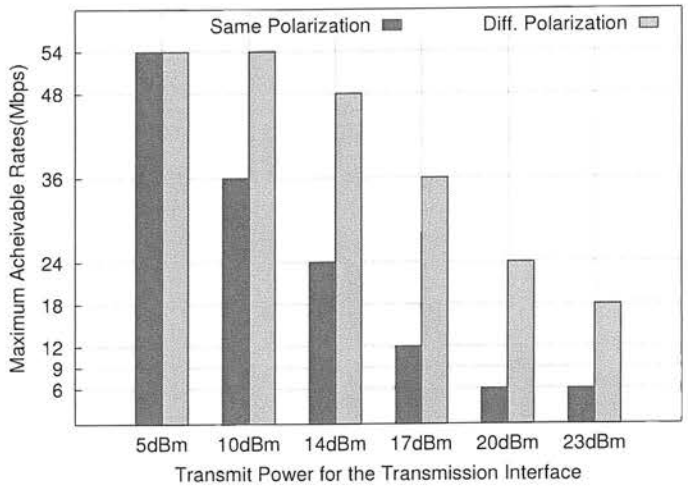


Figure 4.6: Maximum achievable bit-rates on the receive interface at different power levels for the transmit interface when both interfaces are tuned to far apart channels within the 5GHz and their antennas are separated by 40cm.

Figure 4.2 with end nodes using different channels chosen from 36 and 157. After repeating the same experiment as in the previous section to determine the minimum antenna separation for different power levels and bit-rates, we find that *antenna separation has to be greater than 40cm* to avoid performance degradation across all power levels and bit-rates, especially higher bit-rates and power levels (note that in our setup the length of the pigtails connecting the radio interfaces to the antennas limits the antenna separation to maximum 40cm). This is the case even when using differently polarised antennas. However, using differently polarised antennas is still relatively beneficial. For example, the minimum antenna separation requirement in the differently polarised antenna case is five times lower than with identically polarised antennas (6cm vs. 30cm) for the 6Mbps rate and 17dBm transmit power level.

To better capture the benefit from using differently polarised antennas, we provide a comparison in terms of maximum achievable bit-rate at different power levels in Figure 4.6 when the antenna separation is fixed at 40cm. *We observe that using differently polarised antennas* ***permits using bit-rates that are up to 4 times higher*** *for the same interferer power level* (see rates comparison for the 20dBm power level).

4.3.3 Microscopic Observation of Adjacent Channel Effect

Let us consider the case where collocated interfaces on a multi-radio node are assigned *nearby channels*, thereby introducing the additional possibility of *adjacent channel interference (ACI)* besides receiver blocking, which was the only phenomenon causing coexistence interference in our experiments so far. Although ACI characterisation in multi-radio 802.11 networks has received a fair amount of attention in the literature (§2.2.2), it was done indirectly using packet delivery ratio measurements. We aim to complement and validate observations from previous work by taking a direct and microscopic look at MAC behaviour in the presence of ACI and correlate it with packet delivery ratio.

An 802.11 interface can be viewed as being in one of the following four states: transmit (TX), receive (RX), channel busy (BUSY) and IDLE [130]. The BUSY state occurs when energy detected on the channel is more than a specified threshold (*e.g.*, -62dBm for Atheros chipset in 802.11a). In Atheros-based chipsets, three counter registers are updated at 40MHz frequency and show the percentages of time that the MAC is in RX, TX or BUSY states. We access the values of these registers from user-level at 1Hz using shell scripts.

As before, we use the three-node setup shown in Figure 4.2 with Compex cards and Cisco antennas. The separation between antennas on the middle dual-radio node is 40cm. One of the interfaces on that node is receiving data on channel 36 (5.18GHz) sent at 6Mbps bit-rate. The transmit power level of the other interface is set to 17dBm and it shifts channels every 20 seconds from 36 – 40 – 44 – 48 while continuing to broadcast traffic on each of those channels. The result of this experiment is shown in Figure 4.7. The percentage of BUSY time of MAC on the receiving interface of the dual-radio is computed based on tracking register values as described above. The received throughput is also shown in the figure.

Focusing on the case of identically polarised antennas on the dual-radio node, we see that the MAC of the receiving interface finds the channel to be fully busy not just when both collocated interfaces use the same channel (36) but also when the interfering interface uses the immediately adjacent channel (40). The received throughput differs quite widely between these two channels — channel sharing happens between the two interfaces on channel 36, whereas the throughput of the receiving interface drops almost to zero due to the high level of ACI from nearby transmissions on channel 40. As the broadcasting interface moves to channel 44, the receive interface finds the MAC

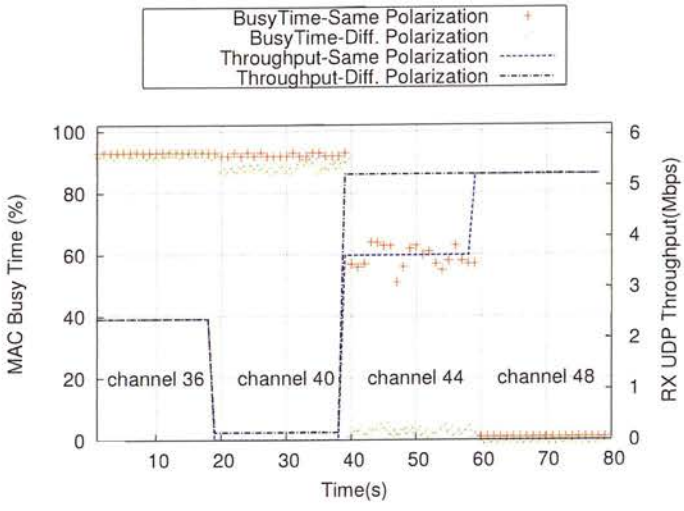


Figure 4.7: The effect of adjacent channel interference (ACI) on receive MAC BUSY time and throughput at different channel separations and polarisations between collocated interfaces.

to be busy only 60% of the time and its throughput increases proportionally. With a separation of three channels (*i.e.*, when the transmitting interface is on channel 48), the effect of ACI on the receive interface MAC busy time and throughput becomes negligible.

Similar qualitative behaviour holds when antennas on the dual-radio are differently polarised except that now only two channel separation is required between the collocated interfaces to avoid performance degradation, once again demonstrating the benefit of using antenna polarisation to mitigate multi-radio coexistence interference.

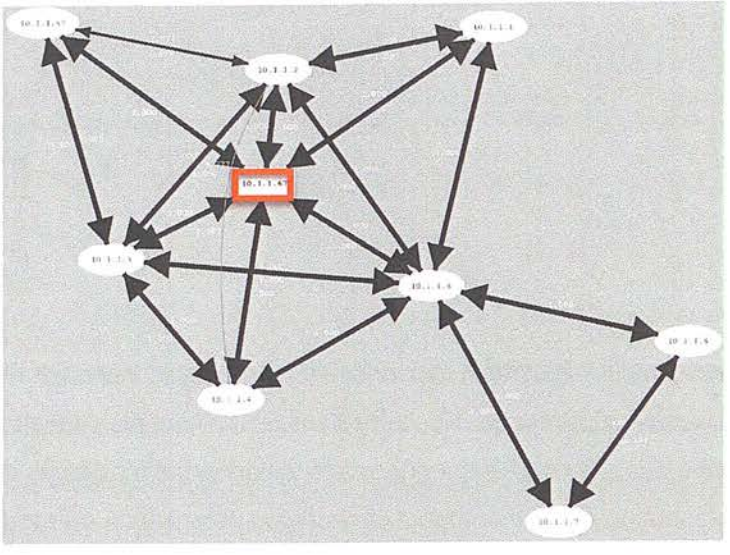
4.4 Effect of Using Different Antenna Polarisations on Network Topology

Our study of multi-radio coexistence interference in both multi-band and single band configurations in the previous two sections showed that using differently polarised antennas on a multi-radio node is helpful in mitigating such interference. However, changing polarisation changes the radiation pattern of the antenna, thus it could potentially have a negative impact on mesh network topology and link qualities. To investigate this issue, we use the indoor multi-radio mesh network testbed deployed in the Informatics Forum building at the University of Edinburgh. The testbed con-

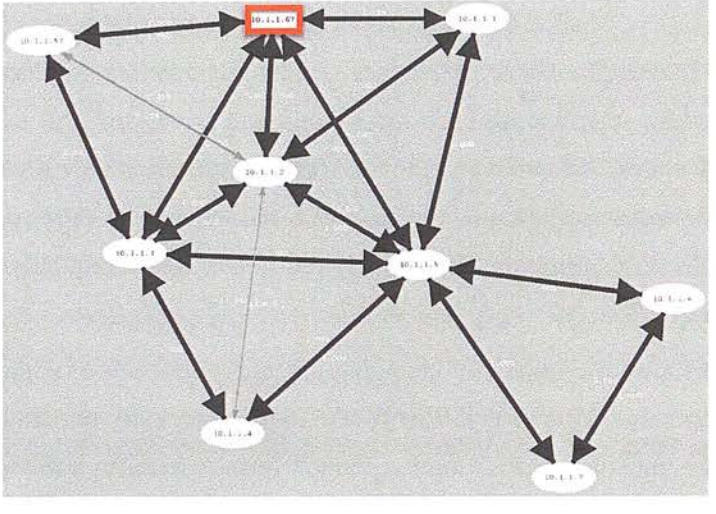
Figure 4.8: Indoor multi-radio mesh network testbed deployed in the Informatics Forum building at University of Edinburgh. Node positions on the floor plan are shown in the figure. Rectangular-shaped node is the one where antenna polarisation is changed from vertical to horizontal to study the impact of changing polarisation on network topology and link qualities.

sists of 9 Avila-based mesh nodes running OpenWrt and equipped with 4 Compex 802.11a/b/g wireless cards each. Positions of testbed nodes on the floor plan is shown in Figure 4.8. In order to isolate the impact of changing antenna polarisation, we only consider one radio interface on each mesh node and configure it to use channel 36 common to all nodes. All nodes run the OLSR mesh routing protocol which relies on the ETX metric to assess link quality. ETX of a link represents the expected number of transmissions on that link and is inversely proportional to the quality of the link in each direction [131].

Figure 4.9 shows the impact of changing antenna polarisation on the rectangular-shaped node on network topology. We observe that connectivity remains unaffected at the node whose antenna polarisation is changed from vertical to horizontal. To get a more detailed understanding of the effect of changing antenna polarisation, Table 4.2 shows various statistics related to topology and link qualities as measured using the ETX metric in OLSR. They confirm that changing antenna polarisation does not significantly alter the topology or link qualities. For example, average node degree drops by around 5% and mean ETX increases by around 20%. We also found that using differently polarised antennas at the end nodes of a link only marginally degrades the throughput (by about 17% on average) compared to using identically polarised antennas. For most links throughput degradation is less than 10%(*i.e.*, Median=8%, stdev= 24%). These results are consistent with earlier modelling and measurement efforts



(a) Same Polarisation



(b) Dif. Polarisation

Figure 4.9: Snapshot of mesh testbed topology when antenna polarisation of rectangular shaped node is changed from vertical to horizontal. Polarisation of radio interface on all other nodes remains fixed at vertical.

Table 4.2: Topology and Link Quality Statistics

	Vertical Polarisation	Horizontal Polarisation
Average Node Degree	4.22	4.00
Node Degree Standard Deviation	1.856	1.803
Mean ETX	1.266	1.538
Median ETX	1.031	1.452
90th Percentile ETX	2.0	2.202

studying the effect of polarisation in indoor environments [122].

Considering the above mentioned benefits of using different polarized antennas, such as increasing up to four times the maximum achievable bit-rates in the single band scenario (Figure 4.3.2)), a 17% average degradation of the throughput performance is a tolerable cost for the benefit.

4.5 Summary

We have experimentally investigated the extent and nature of coexistence interference that occurs between collocated interfaces on multi-radio 802.11 platforms. We show that interfaces with antennas in close proximity on such platforms suffer severe performance degradation regardless of channel separation (including use of different frequency bands). We demonstrated in both single and multi-band configurations and with different channel separations that antenna polarisation can alleviate this effect without adding to expense or causing any serious negative side effects. Moreover, using a direct and microscopic look at MAC behaviour we confirm observations from prior studies on adjacent channel interference characterisation on the need for a minimum amount of channel separation. Our study also reiterates that we cannot do away with antenna and channel separation in the design and configuration of practical multi-radio 802.11 platforms.

Chapter 5

Indoor WiFi Fingerprinting-Based Localisation in Diverse Environments

5.1 Introduction

Localisation or knowing the physical position of an object, thing or person becomes more important with the increasing popularity of handheld devices such as smartphones and tablets. Many mobile apps now require knowing the location of their users in order to provide information based on the location (*e.g.*, shopping aid apps, journey planner apps, social network apps, etc.). GPS[1] is the main technology for localisation — many devices are now equipped with GPS and it is globally available. However, its application in indoor environments, where people spend most of their time, is very limited due to the weakness of signals received from satellites. The popularity of WiFi networks in indoor environments and the existence of WiFi interfaces in almost all smartphones make WLAN location fingerprinting an alternative for GPS in indoor location estimation and also complementary to GPS for outdoor localisation. In location fingerprinting, location is inferred by distinct features, the fingerprint, associated with each location. Specifically, in WiFi fingerprinting, features of receiving WiFi signals from visible APs (*e.g.*, received signal strength) are considered as fingerprints. WiFi fingerprint location estimation has received much attention in the literature with more focus on improving location estimation algorithms and less attention on underlying factors such as definition of fingerprint, effect of WLAN features and their spatio-temporal behaviours (such as differences in the signal propagation between 2.4GHz and 5GHz bands).

[1] Global Positioning System

With respect to the fingerprint definition, eight different definitions are considered that span RSSI based, AP visibility based and combinations of both. They are tried for three different deterministic techniques (including the often used Euclidean distance-based nearest neighbour method) with two probabilistic techniques that use Gaussian and Log-normal distributions for RSSI modelling.

To evaluate the performance of different location estimation algorithms for defined WiFi fingerprints across different locations, a custom mobile application is developed. It performs network scanning by hopping between frequency channels and stores information on visible APs for each location of users pinpointed on the floor map of the building. Captured samples for each position are analysed offline as done in [87].

Results show that deterministic location estimations (*e.g.*, Manhatan, Mahalanobis) outperform probabilistic schemes across different environments. However, the choice of fingerprint definition has as much or more impact than the location estimation algorithm. The best choice of a fingerprint definition and location estimation algorithm is also different between different environments, even between floors of a building.

Two features in WiFi networks are getting more popular: operation in 2.4GHz and 5GHz bands and VAPs. The 5GHz band spatial propagation is different from the usually used 2.4GHz owing to using higher frequency. This motivates us to study how fingerprints received in different bands are different and to see if it can be exploited for better location estimation. Results show that there is a significant benefit in having APs working in 5GHz. More investigations show that the beacons transmitted in the 2.4GHz band are less frequently observed compared to the ones in the 5GHz band during limited numbers of scanning required to build the radio map in the training phase and the online measurement phase. We also observed higher variation in the signal strength from APs operating in the 2.4GHz. Further investigation shows that beacon loss increases in the presence of the co-channel interference which can affect the received signal variation seen in the location estimation system because of having lower number of samples due to beacon loss. This results in higher variation in the signal strength in case of high level of interference, as in an environment with high mobility.

We also consider, for the first time, the effect of virtual access points (VAPs), which are now becoming commonplace in most indoor environments. Our findings show that the presence of VAPs significantly improves WiFi fingerprinting accuracy which we believe is due to two reasons: VAPs have a substantial influence on the AP density, a factor known to affect accuracy with WiFi fingerprinting; and fingerprints obtained

from different co-located VAPs operating on the same channel are somewhat dissimilar, capturing the temporal variability inherent to wireless signal propagation and providing robustness against it.

The findings in this study are summarised as follows:

- Our analysis shows that the fingerprint definition is at least as important as the choice of location estimation algorithm; the latter has received significantly more attention in the literature to date.
- There is no single combination of fingerprint definition and localisation algorithm that always yields the optimum localisation result across all the different environments.
- It is observed that different floors within the same building have different optimum selection of fingerprints and location estimation algorithms.
- Studying the impact of frequency band on WiFi fingerprinting and location estimation shows that operation in 5GHz offers relatively better location accuracy due to lower RSSI variation.
- Contrary to intuition, we find that the presence of VAPs significantly improves WiFi fingerprinting accuracy.

Note that this study is built on the recent study done by Jiwei Li in [87] where the focus was mainly on the impact of different fingerprint definitions and location estimation algorithms on the accuracy of WiFi fingerprinting systems. Specifically, in [87] effects of different components of a fingerprinting location estimation system including fingerprint definition and location estimation algorithms were studied for a multi-storey office building. In this thesis, we follow up that work by studying the importance of fingerprint definition across different environments with different sizes. In [87] — as part of studying various properties of RSSI time series used in constructing fingerprints — the effects of the presence of VAPs are mentioned without studying their impact on the location estimation. In this thesis, we study the effect of VAPs on location estimation for different WiFi fingerprint definitions and explain why the presence of VAPs actually improve location estimation. In addition, we study the effects of multi-band operation on location estimation performance by using mobile phones that support operating in both 2.4GHz and 5GHz bands.

The rest of this chapter is structured as follows. §5.2 explains how WiFi fingerprint information is collected using mobile phones (§5.2.1), introduces different environments for evaluation, including office buildings and shopping centres (§5.2.2) and explains three classes of WiFi fingerprint definitions: RSSI based, AP-visibility based and a hybrid of these in §5.2.3. §5.2.4 introduces location estimation algorithms deployed in this study; they are comprising both deterministic and statistical schemes. In §5.3, the relative importance of fingerprint definition in relation to location estimation algorithms for different environments is studied. §5.4 represents results for the effect of multi-band operation on the location estimation. Effects of the presence of VAPs on WiFi fingerprinting accuracy are studied in §5.5 and a summary of our findings and contributions is given in §5.6.

5.2 Methodology

5.2.1 Data Collection

WiFi fingerprinting data for this study are obtained using Android phones and IndoorScanner, a custom mobile application we developed for this specific purpose. For each measurement position, which we note as the ground truth, IndoorScanner relies on the Android API (specifically, the getScanResults() method in the WifiManager class) to do multiple (20) scans, each taking approximately 1 second. Information gathered from each scan includes service set identification (SSID), basic service set identification (BSSID), RSSI, channel and UNIX timestamp. Scan results are annotated with the corresponding ground truth position and stored in a MySQL database, in a separate table for each different environment. Samsung Galaxy S3 or HTC Nexus One phones are deployed in capturing the WiFi fingerprint dataset. Both smartphones are Android based. Note that part of the datasets for studying the office building are taken from the previous study in [87].

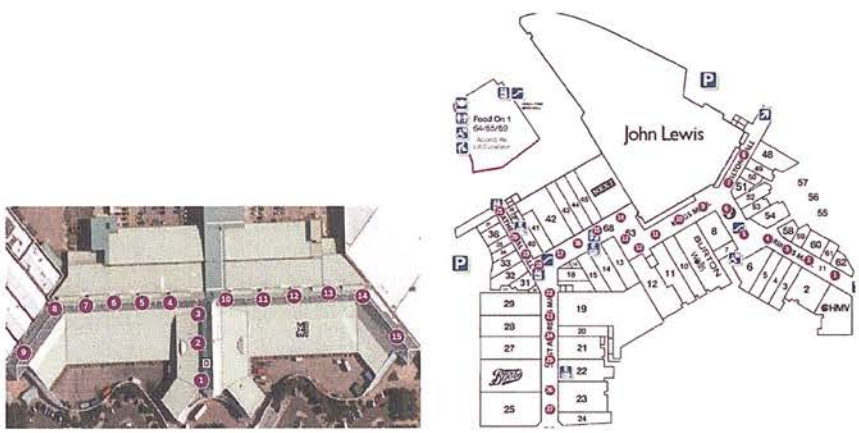
5.2.2 Environments

We consider a multi-storey office building and three different shopping centres as representative set of diverse environments. The layout of these different environments is shown for reference in Figure 5.1 and Figure 5.2.

Multi-storey office building. As a representative office building, we consider the

(a) First Floor (b) Second Floor

Figure 5.1: Floor plans for the first and second floors of Informatics Forum, University of Edinburgh (office environment). Sampled locations during data collection are shown as cyan-coloured cells.



(a) Gyle (Shop. Ctr. 1) (b) St. James (Shop. Ctr. 2)

(c) Ocean Terminal (Shop. Ctr. 3)

Figure 5.2: Layouts of three shopping centres in Edinburgh (shopping centre environments). Purple-coloured cells represent the locations sampled during data collection.

Informatics Forum building in the University of Edinburgh which houses the School of Informatics. We focus on five floors of this building which constitute the main areas with staff/student offices, common spaces and labs. Figure 5.1 shows the floor plan for two of the floors. Note that the grey area in the middle is empty across all floors. Also note that two of the floors, including the second floor shown in Figure 5.1(b), are slightly different with an open plan common space in place of some rooms. As a result the number of sampled measurement locations is different between floors — floors with open spaces have more measurement locations. There is a university-run wireless LAN service across the whole building with several APs installed per floor. Each of these physical APs functions as two virtual APs corresponding to two wireless networks with different user authentication mechanisms. In addition, a number of other APs can be seen across the building, some installed by various research groups in the building while others are located in surrounding buildings. The WiFi fingerprint dataset for this building was generated by measurements using our IndoorScanner app described above along the corridors and in common spaces at a granularity of 1 square meter cells, coloured cyan in Figure 5.1.

Shopping centres. Besides the office building described above, we also consider three shopping centres of different sizes in Edinburgh, UK as shown in Figure 5.2. We use WiFi scan results with our IndoorScanner app along with a distinct ID we manually assigned for each measurement position (shown as purple-coloured cells in Figure 5.2) to produce the individual datasets for each of these environments. Note that compared to the office environment described above, sampling of these shopping environments is sparser as they are public spaces with less flexibility in choosing measurement location and also given their size. These measurements were collected during busy shopping times to better capture a realistic usage scenario.

5.2.3 Fingerprint Definitions

What constitutes a WiFi fingerprint, *i.e.* the fingerprint definition, potentially influences the accuracy of a WiFi fingerprinting system even if other aspects such as the location estimation algorithm are kept fixed.

As it is introduced in §2.5, in WiFi fingerprinting we have a training data set contains all off-line measurements denoted by $\mathbb{T}$ in this study. Subsets of the training dataset (denoted by $\mathbb{S}$) for each cell (measurement location) are chosen based on the fingerprint definition. The online measurement set (denoted by $\mathbb{R}$) used for comparison

is also chosen based on the fingerprint, and the location is estimated using techniques specified in §5.2.4.

In our implementation of the WiFi fingerprint, each cell i has a set of measurements ($\mathbb{T}_i$). Members of $\mathbb{T}_i$ are 5-tuples denoted by $T_{im} = (i, m, \bar{r}_{im}, \sigma_{im}, f_{im})$. In this vector, i is the unique location id representing the cell number; m is another unique identifier for APs seen in the cell i and contains the AP's MAC address; $\bar{r}_{im}$ is the mean of RSSI values measured for beacons received from AP_m in the cell i (note that, in the offline phase several measurements are carried out in each cell); σ_{im} is the standard deviation of the RSSI values for beacons from AP_m in the cell i; finally the number of times AP_m is seen in the cell i is denoted by f_{im} *i.e.* the total number of beacons received through the training phase from that AP. Note that, R_m is defined in the same as T_{im} for the online dataset $\mathbb{R}$, except that i is not known and supposed to be realised by location estimation.

Considering the above representations of parameters for a WiFi fingerprint system, *WiFi fingerprint definition* is defined as *criteria* for choosing T_i tuples from the $\mathbb{T}$ set for each cell i to make a new subset $\mathbb{S}_i$ with length n. The parameter(s) (*e.g.*, mean RSSI $\bar{r}$) that is (are) chosen to be compared with their counterpart in the online data set $\mathbb{R}$ is also defined by the fingerprint definition. We consider eight different definitions that are divided based on the defined criteria.

5.2.3.1 Default

We refer to the WiFi fingerprint definition introduced in [33] as the *Default* WiFi fingerprint definition. *Default* definition does not have any criterion for choosing among the tuples T_{im} in each cell i and simply compares the mean RSSI $\bar{r}$ of all tuples R_m in the online measurement set $\mathbb{R}$ with their correspondents in the $\mathbb{S}$ set.

For the rest of the definitions, the fingerprint also specifies the criterion for choosing n tuples (in our study $n = 5$; a larger set does not show improvement in the location estimation) from the training dataset for each location i. This results in a smaller $\mathbb{S}$ set.

5.2.3.2 RSSI based

Received signal strength (RSSI) of beacons from APs is a key feature commonly considered in WiFi fingerprinting. We consider the following three different fingerprint definitions based on RSSI:

Strength. In this definition, for each location i in the training data set $\mathbb{T}_i$, a subset of

$n = 5$ tuples T_{im} with the *highest* values in the mean of RSSI for all seen APs ($\bar{r}_i$) are chosen to make the subset $\mathbb{S}$.

Stability. This definition focuses on the most stable subset of APs in each cell i based on the standard deviation of their RSSI values (σ_i). So, the first five tuples T_{im} are chosen based on the their minimum σ_{im}. The rationale for considering this definition is two-fold:

- Received signal strength is inherently time varying, the signals that vary less would more likely result in better accuracy of localisation.
- [89] conclude from their analysis that RSSI standard deviation is the most influential factor determining the accuracy of a WiFi fingerprinting system.

Variance: It is based on the observation that seems ideal for a fingerprinting-based localisation system if fingerprints from different cells are sufficiently distinct from each other, *i.e.*, fingerprints serve as unique location signatures. So, for each cell i, n tuples (T_{im}) from subset of $\mathbb{T}_i$ are chosen to satisfy the criterion of highest variance of $\bar{r}_m$ ($\bar{r}_m$ represents a vector of all measured RSSI ($\bar{r}_{im}$) for the AP_m across all cells that it has been seen).

5.2.3.3 AP Visibility based

The visibility of APs is an important aspect for WiFi fingerprinting systems. Some proposals assume that identical sets of APs are seen across the whole space of interest, whereas others implicitly suppose that the visibility of an AP is constant over time. These assumptions often do not hold in practice. To capture the impact of AP visibility on the accuracy of WiFi fingerprinting systems, the two following definitions are considered:

Constancy: At a given location i, f_{im} is defined as the number of times an AP_m is seen across different measurements. Because of weak signals, small-scale fading or beacon loss due to co-channel interference, it happens that beacons from the AP_m is not seen in every WiFi scanning. The constancy definition essentially captures this aspect. Specifically, for each location i in the training data set $\mathbb{T}_i$, a subset of five tuples T_{im} with the *highest* values in f_im are chosen to make the $\mathbb{S}$. The mean RSSI ($\bar{r}_i$) vector of this subset of APs is then chosen for location estimation.

Coverage: This definition captures a different spatial aspect of AP visibility. It picks a subset of five tuples T_{im} from the $\mathbb{T}_i$ with those APs that are *most widely seen across*

all cells. In other words, tuples T_{im} are chosen if their AP_m is visible at a larger more number of cells.

5.2.3.4 Hybrid Definitions

Recall that we have selected a subset of APs satisfying a certain property in our alternative set of fingerprint definitions. However, when using the constancy definition, we observed that often several APs are seen in a cell the same number of times. We randomly break ties with the vanilla constancy definition described above, whereas here we consider hybrid definitions that combine constancy with other similar definitions to further refine the selection. Intuitively, it is more likely to receive more beacons from APs whose RSSI ($\bar{r}$) is higher as they are more prone to beacon loss because of the small scale fading (the same argument is valid for APs with lower RSSI standard deviation); so RSSI fingerprint are chosen to further refine the constancy fingerprint criterion by introducing the two following hybrid metrics:

Constancy+Strength: With this definition, we first rank the APs seen in a cell in the decreasing order of their constancy. Between APs with the same constancy, we prefer those with a higher strength as indicated by their mean RSSI value in the fingerprint database.

Constancy+Stability: As with the previous definition, APs seen in a cell across all measurements in the radio map construction phase are ordered based on their relative constancy so that APs with higher constancy appear earlier in the order. Then stability of the APs as defined above is used to choose among the APs with the same constancy.

5.2.4 Location Estimation Algorithms

In our study, we consider five different location estimation algorithms. The first three belong to the deterministic techniques (*e.g.* RADAR [33]) whereas the other two fall under the category of probabilistic techniques exemplified by Horus [83].

5.2.4.1 Deterministic or Nearest Neighbour (NN) Techniques

The use of nearest neighbour techniques is quite common with WiFi fingerprinting systems. Essentially, the idea is to compute the distance in signal space between pre-collected, location-tagged fingerprints in a database and a runtime fingerprint to find the closest match or matches. Different NN techniques differ in the distance computation methods used. We consider three representative methods as outlined below.

Euclidean Distance: This method used in [33] and other WiFi location fingerprinting systems uses Equation 5.1 to compute the distance between fingerprints from the database, each with an associated location and denoted by $\mathbb{S}$, with a runtime fingerprint $\mathbb{R}$. In Equation 5.1, n is the number of APs considered in the fingerprints; in our study, this is the total number of APs in the environment with the default fingerprint definition and 5 for the other definitions. s_i is the mean RSSI value of AP i in the fingerprint from the database, whereas r_i is AP i's RSSI in the runtime fingerprint.

$$EucDist(\mathbb{S},\mathbb{R}) = \sqrt{\sum_{i=1}^{n}(s_i - r_i)^2} \tag{5.1}$$

Manhatan Distance: Manhatan distance, which is also mentioned in [33], is another well-known NN method. It is defined as the sum of the absolute differences of values between fingerprint from database and runtime fingerprint as indicated by Equation 5.2.

$$ManDist(\mathbb{S},\mathbb{R}) = \sum_{i=1}^{n}|s_i - r_i| \tag{5.2}$$

Mahalanobis Distance: Mahalanobis distance is yet another NN method considered in the WiFi fingerprinting literature (*e.g.* see [89] and references therein). It is more sophisticated compared to the previous two methods and accounts for correlations between compared vectors. An interesting feature of Mahalanobis distance is that it is based on assumptions of stable patterns of RSSI distributions and it also takes into account variance in RSSI as done in probabilistic techniques [132, 133]. Mathematically, Mahalanobis distance computation is shown by Equation 5.3 where S is the covariance matrix of $\mathbb{S}$ and $\mathbb{R}$ of the same distribution.

$$MahalDist(\mathbb{S},\mathbb{R}) = \sqrt{(\mathbb{S}-\mathbb{R})^T S^{-1}(\mathbb{S}-\mathbb{R})} \tag{5.3}$$

5.2.4.2 Probabilistic Techniques

This class of techniques infer the probability that a user is at a certain location based on the probabilistic modelling of RSSI measurements for each cell i from different APs. In simple terms, they estimate the location x of the user that maximises the conditional probability $P(x \mid \mathbb{R})$ given an online fingerprint set $\mathbb{R}$. Different techniques differ in the type of distribution used for RSSI modelling. We focus on two commonly

considered distributions: **Gaussian** (as in [83]) and **Log-normal**. For example, in the Horus system [83] possible locations are estimated as follows:

In the training phase:

- For each access point AP_{im} in each cell (location) i, the histogram of the signal strength is made. This represents the radio map.
- The radio map is approximated by a parametric distributions such as Normal distribution specified by the mean and variance $N(\bar{r}_{im}, \sigma^2_{im})$.

In the online phase:

- Given the online set of signal strength $\mathbb{R}$ with k members (k actually is the total number of seen APs), the goal is finding location x from the possible location sets (denoted by $\mathbb{X}$) to maximise $P(x \mid \mathbb{R})$ (*i.e.*, $\underset{x}{\text{argmax}}[P(x \mid \mathbb{R})]$).
- Using Bayes' theorem:

$$\underset{x}{\text{argmax}}[P(x \mid \mathbb{R})] = \underset{x}{\text{argmax}}[P(\mathbb{R} \mid x)] \tag{5.4}$$

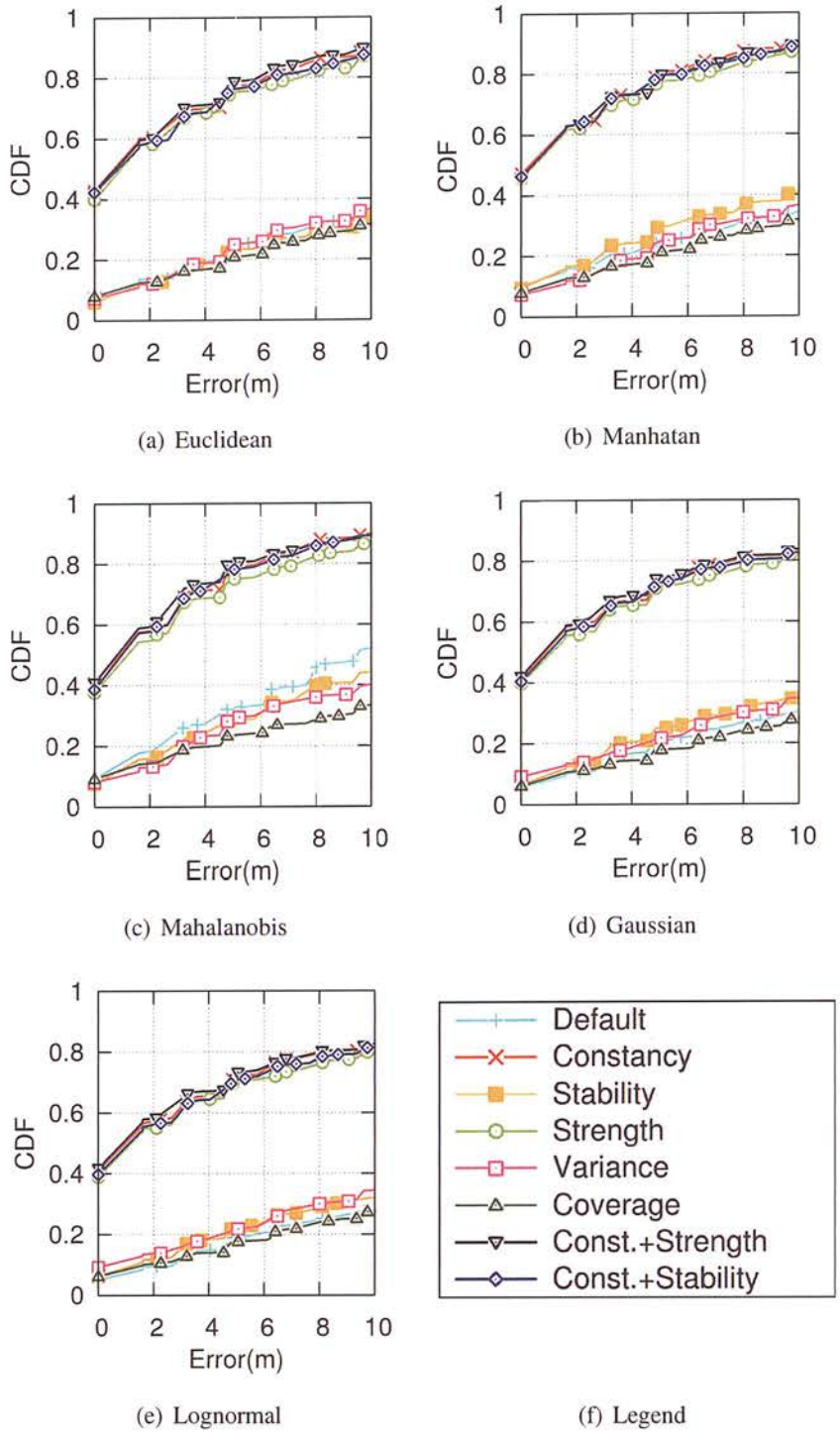
- $P(\mathbb{R} \mid x)$ is calculated using the radio map built in the training phase as:

$$P(\mathbb{R} \mid x) = \prod_{i=1}^{k} P(R_i \mid x) \tag{5.5}$$

5.3 Impact of Fingerprint Definition and Location Estimation Algorithms

In this section, we assess the relative importance of fingerprint definition in relation to location estimation algorithms for different environments. Throughout, we use at least 15 measurement samples (WiFi scans) per location for the reference fingerprint database, and 5 samples for runtime location estimation. We look at the office environment first and then the various shopping centre environments.

Office Environment: Figure 5.3 shows the CDF of location estimation errors with all possible combinations of fingerprint definitions and location estimation algorithms when all 5 floors in the office building are seen as one whole. We see that various fingerprint definitions appear clustered in two separate groups with significant difference in accuracy between them. Constancy, strength and the two hybrid definitions fall in the best performing group. Surprisingly, stability and variance yield poor performance



(a) Euclidean (b) Manhatan

(c) Mahalanobis (d) Gaussian

(e) Lognormal (f) Legend

Figure 5.3: CDF of estimated location errors with different fingerprint definitions and location estimation algorithms **across all floors** in the office environment.

Table 5.1: Office Environment: best combination of fingerprint definition and location estimation algorithm

Floor	Loc. Est. Algo	Fingerprint Defn.
1	Manhatan	Strength
2	Mahalanobis	Constancy+Strength
3	Manhatan	Constancy
4	Manhatan	Constancy+Stability
5	Manhatan	Strength
All	Manhatan	Strength

for all algorithms as does coverage; this is likely due to user mobility and other changes in the environment, that have more effect on those metrics. The Default is also in the same group providing poor location accuracy.

Now turning attention to the various location estimation algorithms, we see that Manhatan distance performs slightly better among the deterministic techniques. It is noteworthy that probabilistic techniques yield poor accuracy compared to all three deterministic techniques; this is more apparent if results are compared near the right end of the plots near 10m error. We believe this is because the true RSSI distribution differs from the one chosen to model it (Gaussian or Log-normal).

Overall, we can also observe that the choice of fingerprint definition has as much or more impact than the location estimation algorithm. Table 5.1 summarises the best combination of fingerprint definition and location estimation algorithm which turns out to be Strength with Manhatan distance for the whole building case.

The best combination is obtained by first identifying the combination providing least median estimation error; in case there are several such combinations then their performance is compared in terms of 3rd quartile estimation errors; if there are still multiple candidates then the one providing the smallest maximum error is chosen as the best combination. It is worth mentioning that the mean of location estimation error over different cells might not be a good metric for finding the best combination as it is very likely that fingerprint definition and location estimation algorithms perform poorly at few cells and their overall good performance should not be biased by high location estimation errors at a few cells.

When each floor is seen in isolation, Table 5.1 also shows that the best combination is different between floors. This is also evident when we look at the median and 3rd quartile estimation errors in Figure 5.4. We see that the second floor has higher errors.

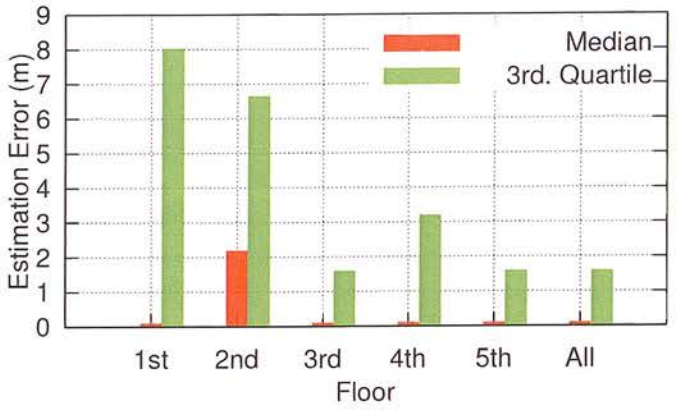


Figure 5.4: Summary statistics (median and 3rd quartile) location estimation error for the best combination of fingerprint definition and location estimation algorithm for various floors separately and together in the office environment.

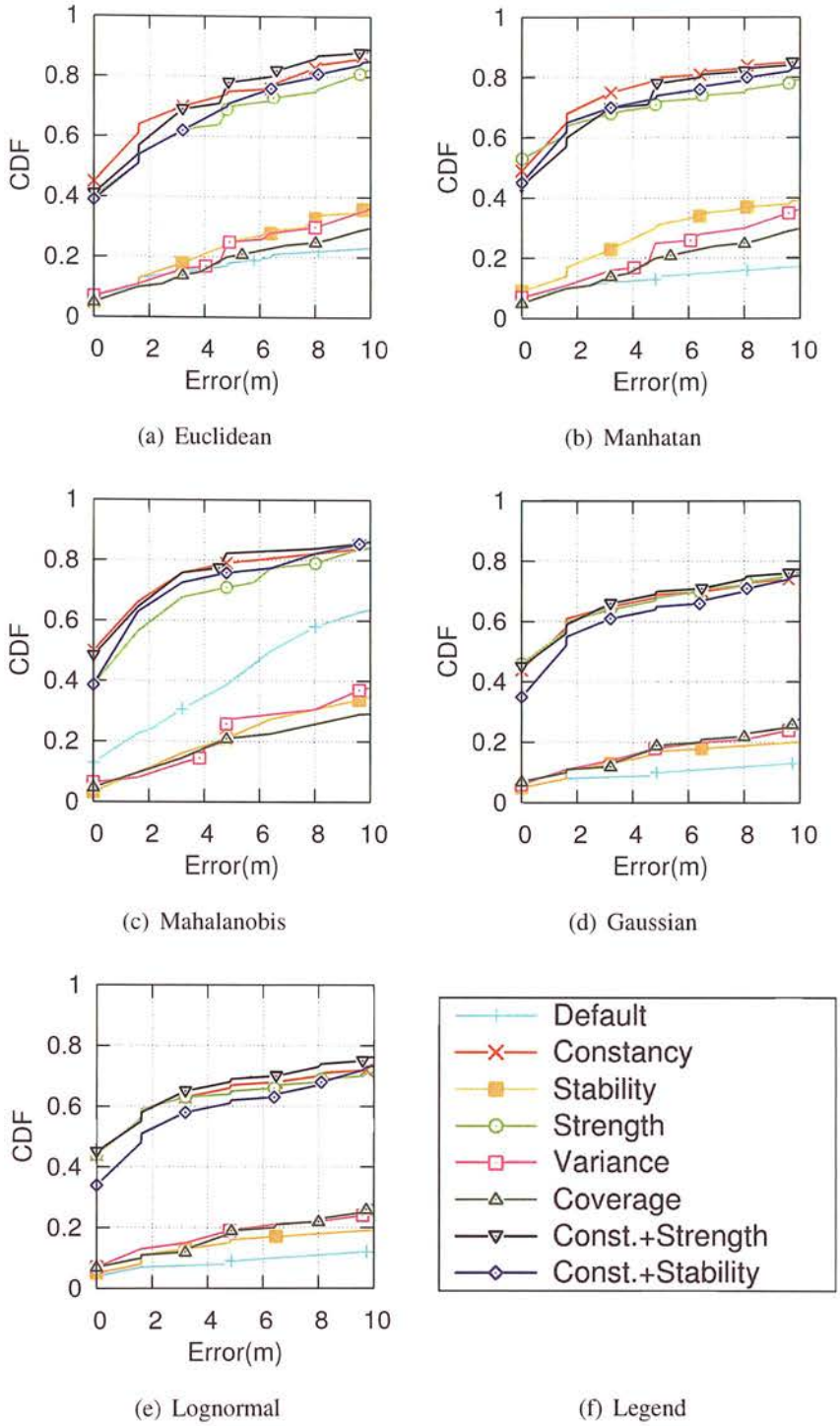
This is because of the open area on that floor where all combinations have difficulty telling apart different cells within that open area. CDFs of location estimation errors for the first and second floors shown in Figure 5.5 and Figure 5.6, respectively, further illustrate this point. We also notice that differences between different fingerprint definitions and location estimation algorithms become more apparent at the individual floor level.

Shopping Centres: Different shopping centres are quite different in terms of their location estimation error statistics as shown in Figure 5.7. We can see that shopping centre 3 is the easier of the three to localise as it is more compact and rich in multipath.

Notice also that errors shown in Figure 5.7 are also higher compared to Figure 5.4, partly because of the sparser location sampling in the former. As with the office environment, we see from Table 5.2 that best combination changes from one environment to the other. This is true even between floors within shopping centre 3, the only one spanning 2 floors in our study. But interestingly, Mahalanobis distance always emerges as the location estimation algorithm in all best combination cases.

5.4 The Impact of Frequency Band

In this section, we explore the impact of the frequency band (2.4GHz vs. 5GHz) on WiFi fingerprinting accuracy. While 2.4GHz was the only band originally used for WiFi, increasingly 5GHz is also being used despite its relatively poorer propagation

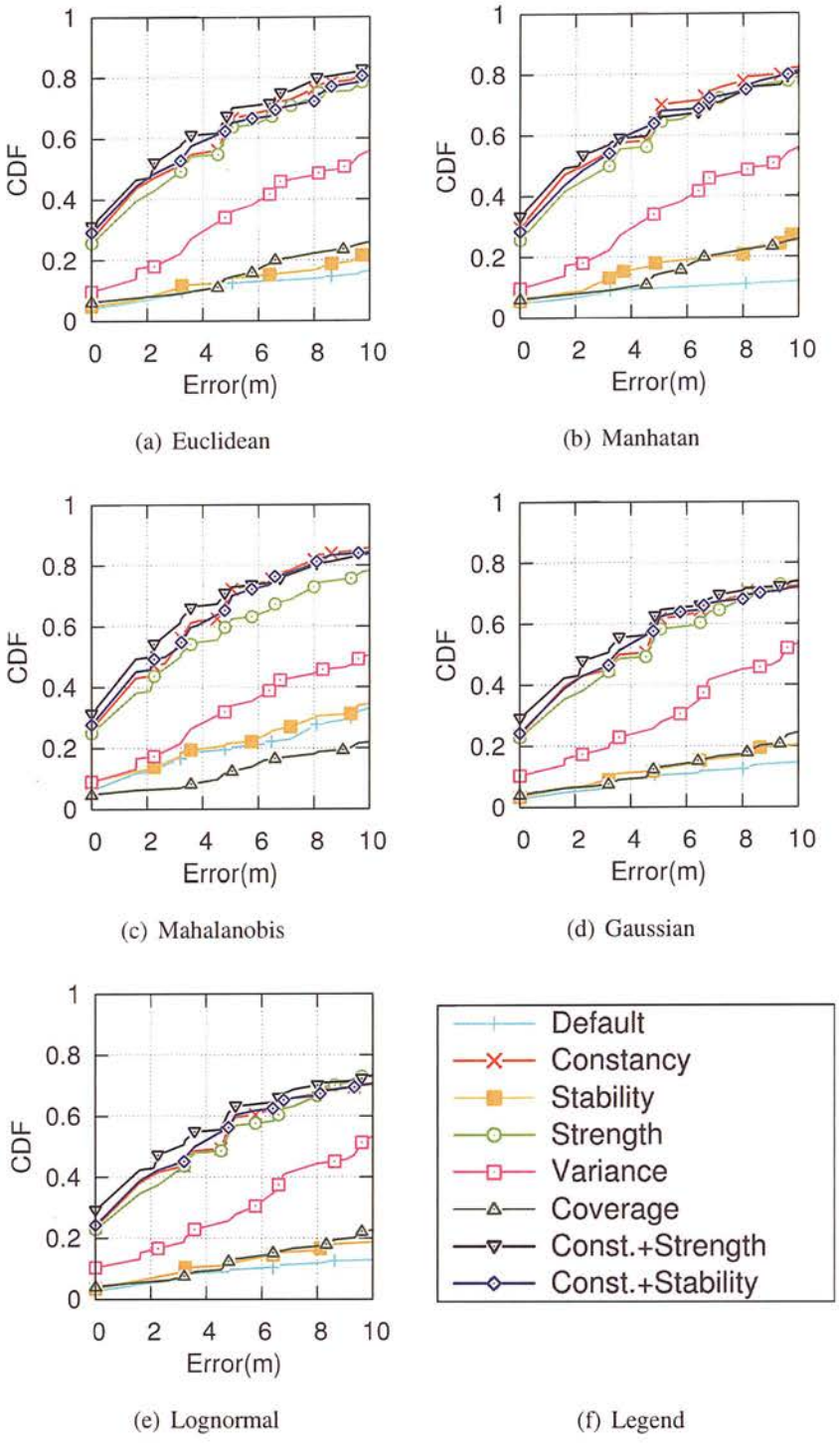


(a) Euclidean (b) Manhatan

(c) Mahalanobis (d) Gaussian

(e) Lognormal (f) Legend

Figure 5.5: CDF of estimated location errors with different fingerprint definitions and location estimation algorithms for **first floor** in the office environment.



(a) Euclidean (b) Manhatan

(c) Mahalanobis (d) Gaussian

(e) Lognormal (f) Legend

Figure 5.6: CDF of estimated location errors with different fingerprint definitions and location estimation algorithms for **second floor** in the office environment.

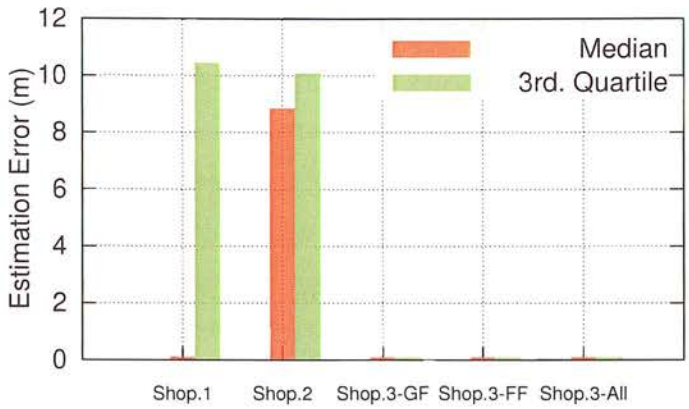
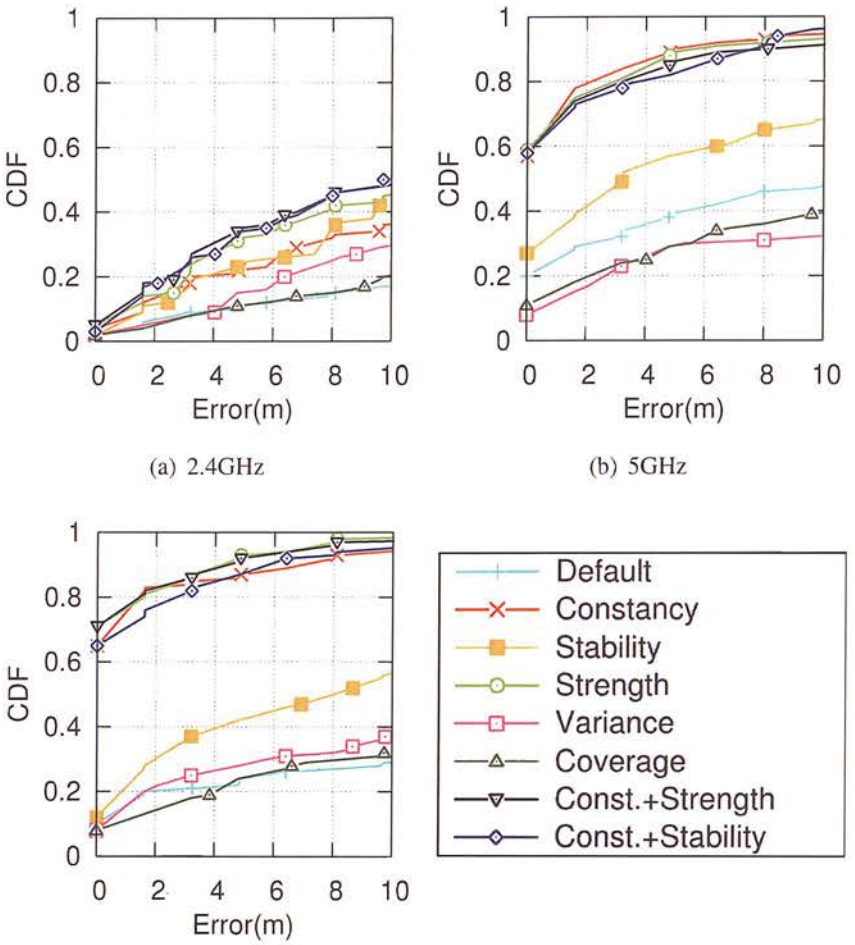


Figure 5.7: Summary statistics (median and 3rd quartile) location estimation error for the best combination of fingerprint definition and location estimation algorithm for different shopping centre environments.

Table 5.2: Shopping centres: best combination of fingerprint definition and location estimation algorithm

Environment	Loc. Est. Algo	Fingerprint Defn.
Shop. Ctr. 1	Mahalanobis	Stability
Shop. Ctr. 2	Mahalanobis	Constancy+Stability
Shop. Ctr. 3-GF	Mahalanobis	Constancy
Shop. Ctr. 3-FF	Mahalanobis	Constancy+Stability
Shop. Ctr. All	Mahalanobis	Constancy



(a) 2.4GHz (b) 5GHz

(c) Both 2.4GHz and 5GHz (d) Legend

Figure 5.8: CDF of estimated location errors across 2.4GHz and 5GHz bands and together for different fingerprint definitions and Euclidean distance method for the first floor in the office environment.

characteristics resulting from higher frequency operation. This is because the 5GHz band is less crowded and also there is far more spectrum available in the 5GHz band. From a WiFi fingerprinting system perspective, in a typical environment today with APs using both 2.4GHz and 5GHz bands, a measurement sample (WiFi scan) obtained either during the training phase or subsequent runtime phase will likely include a mix of 2.4GHz and 5GHz APs. This in turn could impact the accuracy of the WiFi fingerprinting system as signals from these two bands behave differently.

To study the impact of the frequency band on WiFi fingerprinting, we used a smart phone that supports both 2.4GHz and 5GHz bands (Samsung Galaxy S3) to collect multiple samples for each measurement location shown in Figure 5.1(a) for the first floor of the Forum office environment. The dataset is split based on the frequency bands before feeding to the localisation system; that results in having three datasets: (i) the dataset with only 2.4GHz beacons information, (ii) the one with only 5GHz beacons information and (iii) the one that has information for beacons of both bands. Note that, we did not modify the criteria in WiFi fingerprint definitions to be aware of the beacons' received frequency band. In other words, WiFi fingerprint definitions are agnostic to this parameter (defining new WiFi fingerprint that incorporates different bands is left as a future work).

Figure 5.8 shows the CDF of location estimation errors for the cases where only APs from one band are considered as well as the case considering APs from both bands. We show results for only one location estimation algorithm (Euclidean distance) for brevity as the results are qualitatively similar for other algorithms (note that illustrated results are based on the new measurement datasets and are not exactly matched with the one previously shown in Figure 5.5).

Results in Figure 5.8 show that the cases including APs from the 5GHz band show a clear and significant benefit compared to using only the 2.4GHz band even though the number of APs in the environment are evenly distributed across the two bands. Comparing graphs for the 2.4GHz and 5GHz bands also shows that choosing *Constancy* and hybrid fingerprint definitions considerably improves the estimation accuracy. To verify that, we statistically analysed constancy (f_{im}) for each AP (AP_m) across different cells(i).

Table 5.3 tabulates median and mean of the f_{im} for each band. It shows the higher Constancy for APs in the 5GHz band as compared to the ones in the 2.4GHz band while the average received signal strength for all APs from the two bands are very close (see Table 5.4).

Table 5.3: Constancy fingerprint definition for APs in 2.4GHz/5GHz bands for the first floor in the office environment dataset.

	2.4GHz		5GHz	
	Median	Mean	Median	Mean
Constancy	19	15.4	15	13.6

Table 5.4: Mean and standard deviation for the RSSI values received from each AP operating in 2.4GHz or 5GHz bands for the first floor in the office environment.

	2.4GHz		5GHz	
	Mean	Std.dev	Mean	Std.dev
RSSI(dBm)	-82.64	9.47	-83.32	9.74

To better understand the reasons behind the improvement in the Constancy and accuracy obtained using the 5GHz band, we examined how co-channel interference, which is more likely to happen in the 2.4GHz band, might affect it. We setup an AP operating in the 2.4GHz band and had a client in the form of a laptop with an AirPcap USB dongle [60] listening to beacons sent from the AP on the configured channel. We also setup an interfering node (on the same channel) with a modified device driver with CCA (Clear Channel Assessment) disabled so that it can continuously transmit regardless of whether the channel is idle or busy to act as a hidden terminal node. We measured the beacon loss by examining beacons' sequence numbers. The signal strength of beacons are also measured by looking into the PPI (Per Packet Information) headers of frames captured by the AirPcap dongle and Wirshark application.

We find that co-channel interference increases the beacon loss and indirectly affects the variation of the signal strength as explained in the following. Interference can cause beacons loss by acting as a noise in the receiver radio interface and prevent decoding of the frame. Considering the fact that RSSI is measured only for the PLCP header of successfully received frames, interference does not have direct effect on the received signal strength. However, losing beacons turns out as having a fewer samples (lower Constancy) in a limited number of WiFi network scan for both training and measurement phases of WiFi fingerprinting. Having fewer samples in an environment with small-scale fading will result in higher RSSI standard deviation. Table 5.5 confirms that with a high rate of beacon loss (40%) standard deviation of the RSSI is a large number in comparison to when beacon loss is lower (5.94%). However, the mean/median values of the RSSI differ only by 1dB. Note that, as interference has an

Table 5.5: RSSI for receiving beacons from a channel in the 2.4GHz band in presence of co-channel interference.

	without Interference	with Interference
Mean (dBm)	-19.56	-21.85
Median (dBm)	-20	-21
Stdev	0.69	3.1
Beacon Loss(%)	5.94%	40.0%

indirect effect on the RSSI variation, it is unclear if the level of interference causes only few beacon losses as in the case of the measurement reported in Table 5.4. However, it has a direct effect on the Constancy fingerprint definition.

5.5 The Effect of Virtual Access Points

In this section, the effect of virtual access points (VAPs) on WiFi fingerprinting accuracy is studied. VAP is a way to realise multiple APs, each potentially using a different security mechanism and targeting a different set of users, with a single physical AP via time sharing. It is the wireless counterpart of VLANs. The BSSIDs of VAPs corresponding to a physical AP are typically derived from the BSSID (MAC address) of the physical AP. From our study of WLAN deployments in offices and public spaces, we observe that VAPs are common today (more information is provided in §2.1.2).

Our interest here is to study the impact of the presence/absence of VAPs on WiFi fingerprinting. Towards this end, we studied the effects of VAP presence for the first floor of the office environment. As noted earlier, each of the physical APs in the university WLAN network advertise two VAPs. On the first floor there are 33 university-run APs resulting in 66 VAPs, plus 10 other non-VAP APs. Thus in total there are 76 APs when VAPs are counted, and 43 otherwise. In this environment we find that BSSIDs of VAPs share the first ten digits with the BSSID of their corresponding physical AP. It is relevant for WiFi fingerprinting to understand how the beacons of VAPs are transmitted. By capturing all beacons in the air with a laptop running the Kismet application, we find that beacons for each of the VAPs corresponding to a physical AP are sent within a short period of 100ms, the default beacon transmission interval. This suggests all VAPs can usually be detected via passive scanning as the time spent on a channel before hopping to another channel is 100ms by default.

To study the effect of VAPs, we consider two cases, one with VAPs included and the

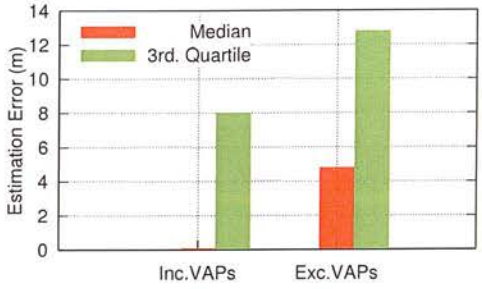


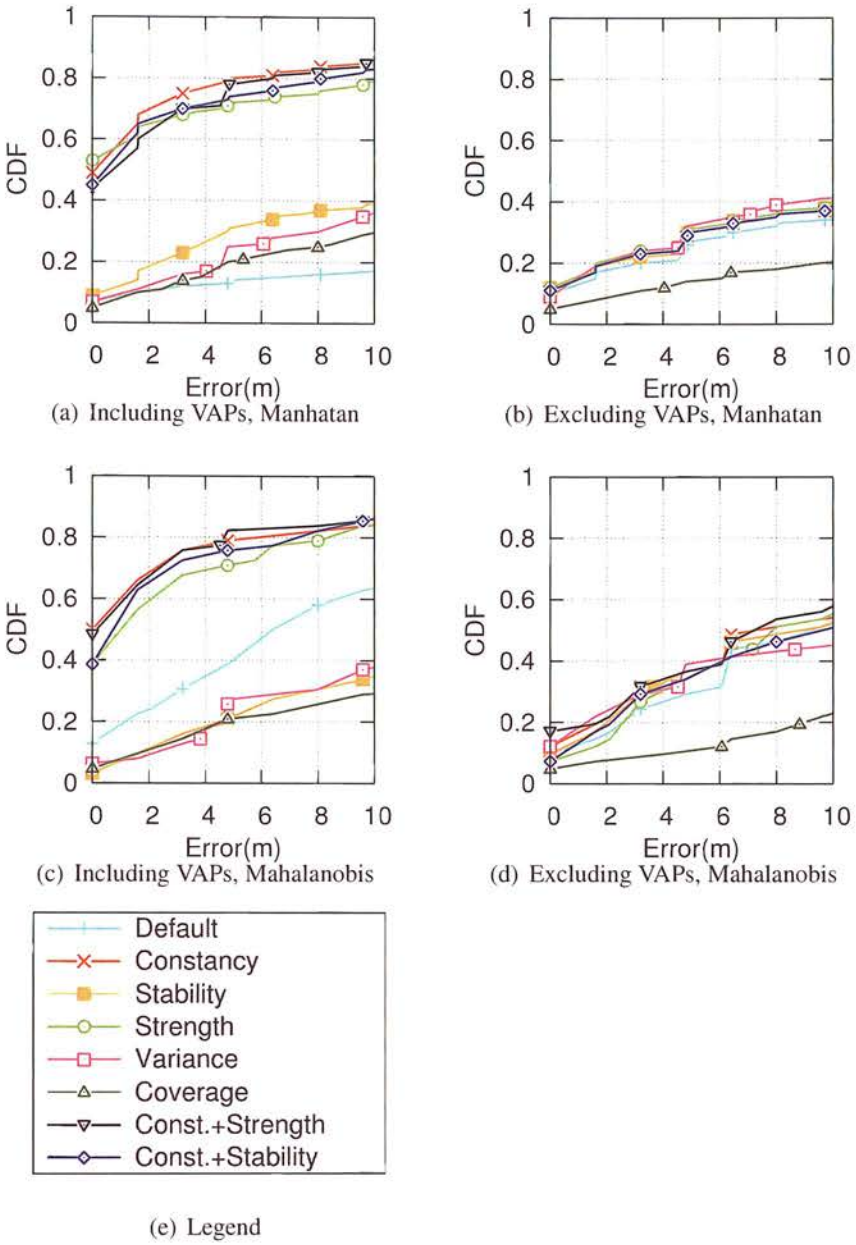
Figure 5.9: Summary statistics (median and 3rd quartile) of the location estimation error for the best combination of fingerprint definition and location estimation algorithm with VAPs included and excluded for the first floor of the office environment.

other in which VAPs are excluded. The case with VAPs included simply treats each VAP as a separate physical AP; this is what we did so far in this study. In contrast, only one VAP per physical AP is retained in the latter case. Figure 5.9 differentiates between these two cases in terms of their median and 3rd quartile errors considering the best combination of fingerprint definition and location estimation algorithm for each case (see Table 5.6). Clearly, including VAPs significantly reduces location estimation error, especially in terms of median. Figure 5.10 demonstrates the benefit from considering VAPs in more detail.

Table 5.6: Best combination of location estimation algorithm and fingerprint definition including and excluding VAPs.

Case	Loc. Est. Algo	Fingerprint Defn.
Including VAPs	Manhatan	Strength
Excluding VAPs	Mahalanobis	Constancy+Strength

We attribute the gain seen from including VAPs to two reasons. Firstly, including VAPs increases the AP density which tends to have a positive correlation with higher location accuracy for WiFi fingerprinting systems. For the results shown here, the case including VAPs has 76 APs in total whereas excluding VAPs brings that down to 43, both for the same area. Secondly, even though we may expect VAPs corresponding to a physical AP to have identical signal characteristics, this is not always the case as beacons from different VAPs are separated in time (*i.e.*, 25ms in the office building), each capturing a slightly different time-varying environment context as demonstrated by Figure 5.11 and Figure 5.12.



(a) Including VAPs, Manhatan

(b) Excluding VAPs, Manhatan

(c) Including VAPs, Mahalanobis

(d) Excluding VAPs, Mahalanobis

(e) Legend

Figure 5.10: CDF of estimated location errors including and excluding VAPs for different fingerprint definitions and Manhatan/Mahalanobis distance methods for first floor in the office environment.

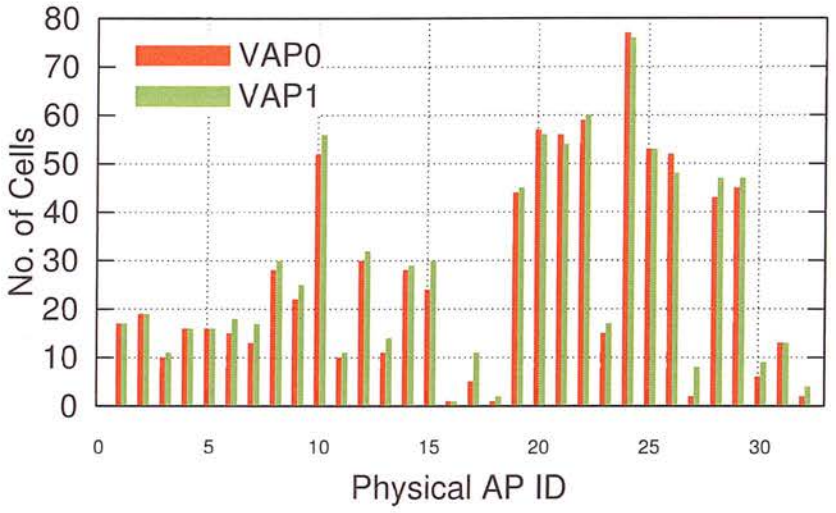


Figure 5.11: Relative differences in signal coverage between each pair of VAPs corresponding to a physical AP in terms of cells where they have been seen.

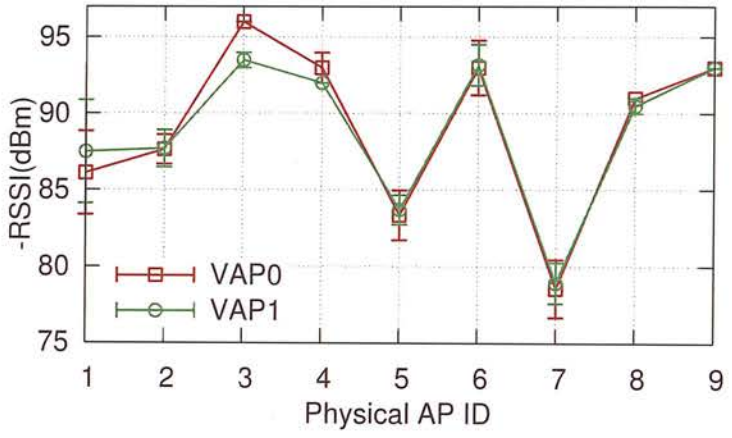


Figure 5.12: Differences in mean and standard deviation of RSSI of each pair of VAPs as seen from a cell that shows maximum improvement in location accuracy from including VAPs.

5.6 Summary

The impact of fingerprint definitions along with location estimation algorithms on WiFi fingerprinting location accuracy across diverse environments has been studied. Results show that the combination of fingerprint definition and location estimation algorithm that yields the best location accuracy is highly dependent on the environment and even on the specific floor within a given environment. It has been shown that the choice of frequency band (2.4GHz vs. 5GHz) and inclusion of VAPs has a significant impact on the location accuracy of WiFi fingerprinting systems, and potential reasons were explained for these findings.

Chapter 6

Available Bandwidth Estimation over Next Generation WiFi Networks

6.1 Introduction

End-to-end available bandwidth estimation (ABE) has a wide range of uses from adaptive application content delivery, transport-level transmission rate adaptation and admission control to traffic engineering and peer node selection in peer-to-peer/overlay networks [1, 2]. For instance, adaptive media streaming services [34, 35, 36, 37] keep multiple versions of each video with different encoding bitrates and stream the content to a client with a bitrate that closely matches the available bandwidth along the end-to-end path to the client. Media servers therefore typically employ bandwidth estimation techniques (*e.g.*, packet-pair) within their streaming protocols [134] to accurately estimate available bandwidth for adaptive streaming and enhanced user experience.

Given it is increasingly popular to access the Internet using the WiFi with an increasing number of WiFi-enabled devices such as smartphones, tablets, laptops, smartTVs, etc. possible effects of the underlying technology, 802.11-based communication, on the performance of the end-to-end available bandwidth estimation become an important issue. The WiFi technology based on the 802.11 suite of standards has evolved significantly in the past decade and a half in view of its widespread use and the growing demand for wireless speeds to match wired Ethernet. The current in the 802.11 series of standards is 802.11n [49], which can provide throughput above the MAC layer nearly reaching 400 Mbps. The follow-on standards (*e.g.*, 802.11ac [135]) promise gigabit wireless speeds. New features in both PHY and MAC layers contribute to improving the link performance including MIMO features and channel bonding in

physical layer and frame aggregation and block acknowledgement for a more efficient MAC layer.

Among new features introduced in the 802.11n standard, in this study we demonstrate that FA has the highest impact on the performance of the ABE. Frame aggregation, as the name suggests, aggregates several frames together and reduces the protocol overhead (*e.g.*, headers, inter-frame spaces, backoff) over the set of aggregated frames, thereby significantly improving MAC protocol efficiency. It is shown in other studies (*e.g.*, [18]) that enabling FA can considerably improve the traffic performance. Given the importance of FA, our main goal is to study the impact of it on end-to-end available bandwidth estimation (ABE). ABE typically involves the use of active measurements with probing packets; the flow of probing packets can be affected by the use of frame aggregation. Previous works proposed for ABE in the context of WiFi networks (*e.g.*, [38, 39, 40, 41, 42, 43, 44]) only consider the effects of the legacy WiFi networks based on the IEEE802.11a/b/g. For the first time, we study the effects of the new features of fast replacing WiFi technologies (*e.g.*, IEEE802.11n) on ABE.

Towards this end, a measurement-based study is conducted using an indoor 802.11n wireless testbed. Three WBest [40], DietTopp [39] and pathChirp [46] tools are candidates for this study given they are representing different ABE estimation approaches and have previously been evaluated in the legacy WiFi networks. We study their available bandwidth estimation accuracy in the presence of 802.11n frame aggregation across a wide range of cross-traffic scenarios. Thus, through this study we make the following contributions and findings:

- Available bandwidth estimation is considered for the current generation of WiFi technology based on the 802.11n for the first time, and it shows that frame aggregation is the most dominant feature among 802.11n features affecting ABE. The same findings can also be applied for the new 802.11ac standard as frame aggregation is specified as a mandatory feature and behaves similarly as in the 802.11n standard.

- Comparison of different ABE tools in various cross-traffic scenarios focusing on the impact of FA leads to the following observations:

    1. The FA feature significantly hurts the accuracy of all ABE tools considered.
    2. DietTopp and pathChirp, the tools that follow the Probe Rate Model (PRM) with different probing approaches, are relatively more robust in the pres-

ence of FA compared to WBest that belongs to the Probe Gap Model (PGM) category.

3. The tools tend to underestimate the available bandwidth in the presence of FA across many different cross-traffic rates in all scenarios.

- Keeping in mind that PGM-based tools like WBest are better suited for the fast ABE needed for adaptive multimedia streaming services, our motivating use case, we take a deeper look at the FA effect on the working of WBest and come up with the two principles of jumbo probes and larger number of probes that together make up our solution approach for improved ABE in the presence of FA. We proceed to develop an enhanced variant of WBest which incorporates our approach and show that it is indeed effective in achieving robust and accurate ABE over FA-enabled 802.11n networks. Although our approach is evaluated only in the context of WBest, we believe that it is more widely applicable.

The necessary background and more information about subjects discussed in this chapter can be found in §2.1.3 and §2.6 from Chapter 2. The rest of this chapter is structured as follows. In §6.2, the experimental methodology including details on testbed and different cross-traffic scenarios is explained. Results in §6.3 justify why FA is chosen among other introduced features in 802.11n. §6.4 presents results and analysis of the impact of the FA on the ABE tools under different cross-traffic/interference scenarios. A new FA-aware tool, WBest+, is proposed in §6.5. Finally, findings and contributions are summarised in §6.6.

6.2 Methodology

Frame aggregation can affect the available bandwidth estimation (*e.g.*, by causing multiple probe packets to get packed inside a single frame). Our broad aim is to characterise the impact of FA on end-to-end ABE in various 802.11n wireless LAN (WLAN) scenarios. We limit our focus to the *active measurement* approach. While passive measurement of available bandwidth or utilisation may be appropriate for the last hop path segment with an 802.11 wireless link, it requires cooperation of intermediate access points (APs) for end-to-end ABE, making it less practical than the active measurement approach. Moreover, passive techniques usually rely on lower-level information from the system and device drivers which may require superuser privileges (*e.g.*, for packet captures) or device driver manipulation. For this study, we choose three representative

active measurement tools for ABE (as already noted in the previous section): WBest, DietTopp and pathChirp. The selection of these specific tools is based on the fact that all three of them are publicly available and on several additional considerations.

WBest has been specifically designed for 802.11 wireless LANs and represents the PGM class of ABE tools. DietTopp, on the other hand, falls into the PRM class of tools and has been considered in various evaluation studies of ABE over 802.11 wireless LANs (*e.g.*, [39, 41]). We include pathChirp, which also belongs to the PRM category, as it has been shown to yield good results in some ABE evaluation studies in wireless network settings (*e.g.*, [107, 42]). Moreover, DietTopp and pathChirp, while from the same PRM class, differ in their probing traffic pattern, thereby allowing us to understand the relative effectiveness of different patterns.

Testbed. We take an experimental approach using an indoor 802.11n wireless LAN testbed (illustrated in Figure 6.1) that emulates typical 802.11 WLAN deployments in home and hotspot environments in a simplified form. It consists of two co-located WiFi networks—the available bandwidth estimation occurs from a server to a client (node A) in the left network while the right network causes interfering traffic. In the left network, two 802.11n clients (nodes A and B) are connected to 802.11n enabled access point $AP1$ over a real wireless channel and $AP1$ is connected to a server S via a Gigabit Ethernet link. The Gigabit link is chosen to reflect the recent increased availability of high-speed broadband services [136] that support around 100Mbps, which renders even 802.11n wireless access links bandwidth-limited in our setting (wireless link capacity is 60Mbps with FA only enabled). Because any of the 100Mbps or 1Gbps links makes wireless access links to be the bottleneck in our setting, we simply use a 1Gbps link. In the right network, node C is connected to another server via access point $AP2$.

All the 802.11n hardware in our testbed is based on Atheros chipsets and is used via the ath9k wireless driver [137]. To avoid external interference from other operational WiFi networks in the surrounding environment of our testbed, we set both $AP1$ and $AP2$ (and as a consequence for all client nodes A, B and C as well) to operate on channel 149 from the 5GHz band which was identified as being unused by other surrounding networks.

The modulation and coding scheme (MCS) determines the maximum data rate. MCS is automatically controlled by the ath9K wireless driver based on the chosen MCS index for transmissions [137]. MCS index 7 is chosen in our experiments as it supports single stream and the maximum possible number of aggregated frames

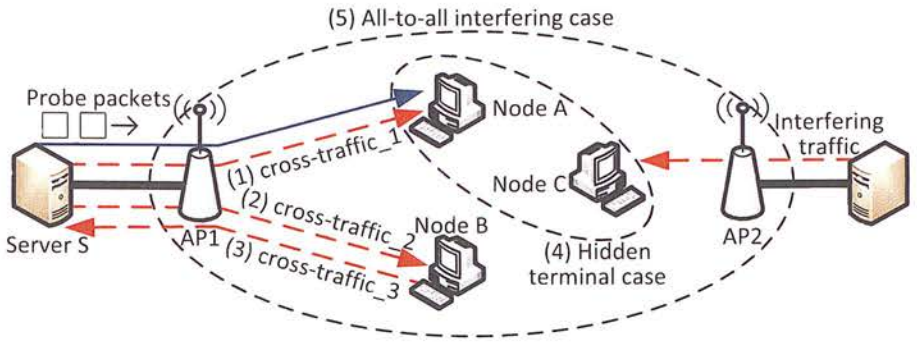


Figure 6.1: Schematic of testbed and various cross-traffic scenarios.

in comparison with other MCS indexes. FA is simply disabled by configuring the maximum possible number of aggregated frames to one frame; managing FA behaviour is added by hacking the ath9k device driver.

Cross-traffic scenarios. As the extent of FA influence on ABE potentially depends on the nature of cross traffic, we consider a wide range of cross-traffic scenarios, reflecting some of the key types of cross traffic that would occur in practice. These scenarios are illustrated in Figure 6.1 and described below. The first three scenarios model various types of cross traffic within a single AP WLAN setting, whereas the last two model cross traffic due to interference from a co-located WLAN. Probing measurement traffic is considered in the downstream direction to the WLAN client (node *A*) for all scenarios to reflect a case where a multimedia streaming server wants to determine the bandwidth available to a user. The level/amount of cross traffic in each scenario is a variable parameter. In all scenarios, cross traffic is generated as a UDP flow using the well-known Iperf tool with a default packet size of 1470 bytes and a specified generation rate to realise different levels of cross traffic.

- ***Scenario 1:*** *Single node case.* In this scenario, there are two flows destined to node *A* from server *S*; one is a probing measurement flow using one of the three ABE tools considered (WBest, DietTopp or pathChirp) and the other is a cross-traffic flow (cross traffic1 in Figure 6.1). This scenario models cross traffic that reflects other downstream application traffic such as P2P file download.

- ***Scenario 2:*** *Cross traffic to Node B.* Different from Scenario 1, this scenario models a situation where another user (node *B*) within the WLAN (*e.g.*, a home WiFi network) competes with node *A* for network bandwidth, for instance, via file downloading or web browsing application traffic. The only source of cross

traffic in this scenario is shown as cross-traffic_2 in Figure 6.1 that competes with the measurement probe traffic (shown with a blow arrow) to get access to the channel.

- ***Scenario 3:*** *Cross traffic from Node B.* The main difference between Scenario 2 and this scenario is in the direction of cross traffic. $AP1$ arbitrates channel access to both measurement and cross-traffic clients in Scenario 2, while B and $AP1$ contend for channel access in Scenario 3.

The following two cases explain how we generate interfering traffic. These cases are tested to answer how tools' estimation performance is affected when there is an interfering channel. These cases are very likely to occur in practice because it is very common that multiple access points are deployed and operate in an uncoordinated fashion as would be the case in a residential area. More specifically, we focus on two interference cases.

- ***Scenario 4:*** *Hidden terminal case.* In this scenario, $AP1$ and $AP2$ are outside the communication range of each other but their associated client nodes (A and C) are close enough to hear each other and to hear transmissions from both APs.

- ***Scenario 5:*** *All-to-all interference case.* In this scenario, all nodes (access points $AP1$ and $AP2$ and their associated clients A and C) hear from and talk to each other. This scenario models a commonly occurring situation with neighbouring home WiFi networks with overlapping coverage areas. Note that in Scenarios 4 and 5, the only source of cross traffic is from the other server to node C via $AP2$.

Note that in Scenarios 4 and 5, we deliberately do not generate cross-traffic traversing on access point $AP1$. Therefore, the available bandwidth estimated by tools in Scenarios 4 and 5 is actually equivalent to a capacity estimate which can vary by the amount of interfering traffic.

In the evaluation of the considered tools under the different scenarios mentioned, we execute 20 experiments with each ABE tool and 95th percentile confidence interval of ABE values is used to show the range of values lie within two standard deviations of the mean.

Obtaining ground truth. For each of the cross-traffic scenarios, we need *true* available bandwidth to assess the accuracy of different ABE tools under consideration (WBest, DietTopp and pathChirp). For this, we follow an approach similar to that

taken in [44]. For Scenarios 1–3, we use backlogged Iperf UDP flow (for 100 seconds) on the probing measurement flow path to find out the true capacity and subtract the level of cross traffic injected to obtain the true available bandwidth. In Scenarios 4 and 5 with interference traffic from another WLAN, the actual available bandwidth is computed as the throughput obtained for a backlogged Iperf UDP flow on the probing measurement flow path while cross traffic on the interfering path is present at the specified level.

Metrics. Our primary focus, as more justified in the next section, is on studying the accuracy of different ABE tools in the presence of FA in different cross-traffic scenarios and levels of cross traffic. The two following metrics are considered to quantify the accuracy of a tool:

i)*Absolute Error:* is defined as the absolute difference between the estimated and true available bandwidth values as specified in Equation 6.1.

$$AbsoluteError = |AB_{estimated} - AB_{true}| \tag{6.1}$$

ii)*Relative Error:* is computed as specified in Equation 6.2.

$$RelavtiveError = \frac{AB_{estimated} - AB_{true}}{AB_{true}} \tag{6.2}$$

We also touch upon two other metrics commonly considered when evaluating an ABE tool, namely measurement duration/latency and intrusiveness (measurement overhead). Note that the method used for obtaining the true available bandwidth (*i.e.*, via a backlogged Iperf UDP flow) is not suitable for ABE in practice because it is highly intrusive in terms of overhead compared to ABE tools with carefully chosen probing traffic (packet pairs, trains or chirps).

6.3 Importance of Frame Aggregation Relative to Other 802.11n Features

Before conducting an in-depth measurement study, in the context of available bandwidth estimation we first elaborate on the rationale behind the choice to focus on frame aggregation among the various new features introduced as part of 802.11n that also include channel bonding (CB) and spatial division multiplexing (SDM). Note that other new features like SGI (Short Guard Interval) and STBC (Space Time Block Coding) are not considered in this study as enabling SGI is only beneficial if the link quality is

very good and it is usually turned off. STBC is designed to increase the reliability of the link and is not tried in this study.

Figure 6.2 shows absolute errors for the different ABE tools used when each of these three features is disabled relative to the case where all of them are enabled. Clearly, the absence of FA leads to the smallest ABE error for all tools. In other words, enabling FA results in the biggest increase in estimation error, in comparison with SDM and CB.

The importance of frame aggregation for achieving high application layer throughput performance is also studied by measuring maximum UDP and TCP throughputs using Iperf for a 802.11n link in our testbed. Figure 6.3 shows achieved throughput for both UDP and TCP traffics on a 802.11n wireless link with and without FA while keeping all other features including SDM and CB enabled. It clearly shows that turning FA on results in significant increase in throughput of both types of traffic. This result reiterates the importance of FA in 802.11n networks.

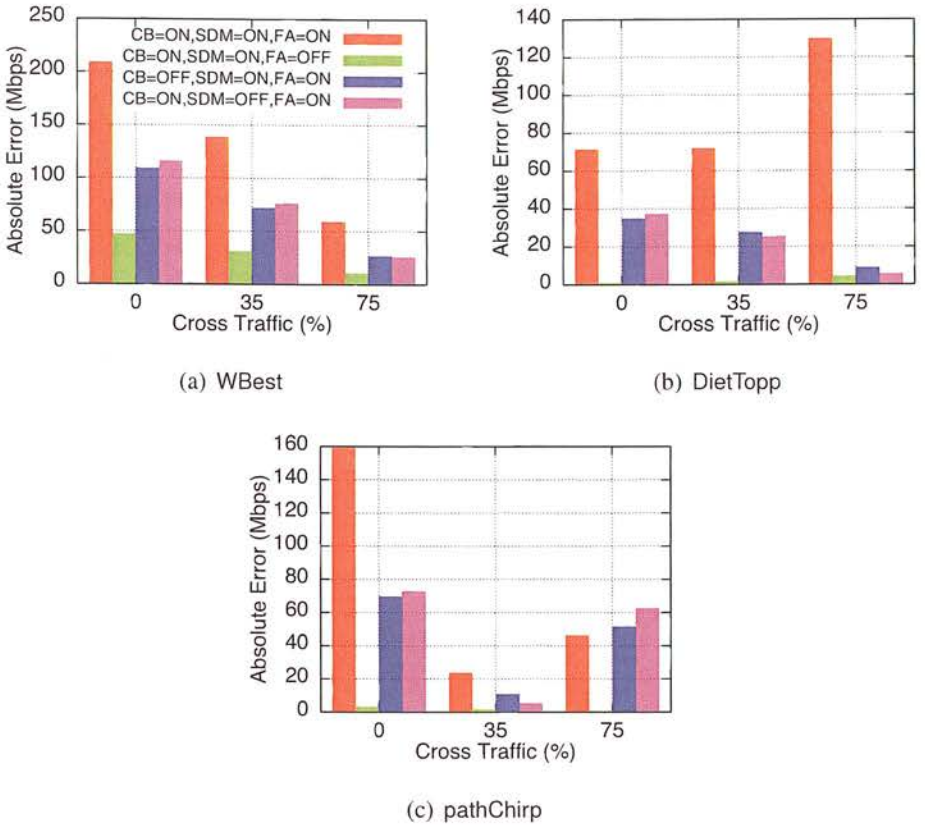
Therefore, FA is chosen among other different features to study its impact on the performance of ABE tools. To isolate the effect of FA on ABE, we disable other 802.11n features including SDM and CB and leave the characterisation of the impact of those other features on ABE for future work. For the same reason, we factor out the link quality- and rate diversity-related effects by ensuring all links operate reliably at the maximum bit-rate possible without SDM (*i.e.*, MCS index 7 in 802.11n which corresponds to 65Mbps physical layer bit-rate).

6.4 Effect of Frame Aggregation on Performance of Available Bandwidth Estimation Tools

In this section, we compare different ABE tools (WBest, DietTopp and pathChirp) in various cross-traffic scenarios focusing on the impact of FA on their accuracy.

6.4.1 Performance Evaluation in a Contention-Free Scenario

In Scenario 1, contention and interference among multiple links is absent, so FA can have a more pronounced effect. Results shown in Figure 6.4 are for evaluating the effect of FA on the accuracy of WBest, DietTopp and pathChirp tools — all other 802.11n features are disabled. The amount of cross traffic is varied and shown as percentage values on the x-axis with respect to the path capacity. For 100% cross-traffic rate,



(a) WBest

(b) DietTopp

(c) pathChirp

Figure 6.2: Relative impact of key 802.11n features (FA, channel bonding (CB) and spatial division multiplexing (SDM)) on the accuracy of different ABE tools in Scenario 1 (see Figure 6.1).

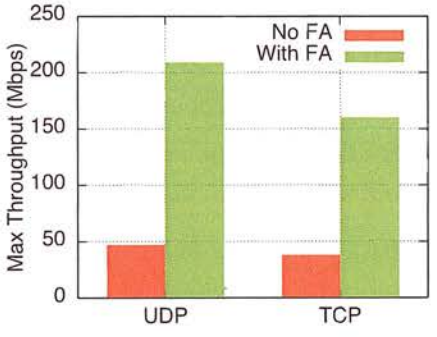
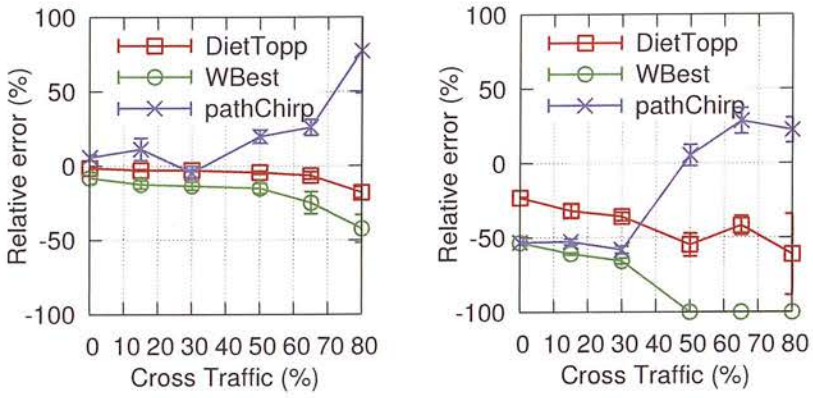


Figure 6.3: Maximum throughput with(out) FA while having all other 802.11n features such as channel bonding (CB) and spatial division multiplexing (SDM) enabled.



(a) FA OFF (Capacity = 30Mbps)

(b) FA ON (Capacity = 60Mbps)

Figure 6.4: Accuracy of DietTopp,WBest and pathChirp in Scenario 1 at varying levels of cross traffic. The median relative error across all cross-traffic levels for DietTopp, WBest and pathChirp are 4%, 14% and 16% respectively when FA is OFF; and 39%, 83% and 24% when FA is ON.

some of the tools under test did not converge so results are only reported up to 80% cross-traffic rate.

ABE accuracy when FA is OFF. Figure 6.4(a) illustrates that DietTopp, WBest and pathChirp achieve good ABE accuracy across most cross-traffic rates; DietTopp performs better in estimation than WBest and pathChirp. DietTopp suffers at most 18% error for 80% cross-traffic rate (*i.e.*, 80% of capacity) while WBest has 43% error at the same rate. DietTopp and WBest tend to underestimate ABE. pathChirp, on the other hand, exhibits overestimation, which coincides with the conclusion in [108]. Also note that the relative error metric somewhat amplifies the error values with increasing cross-traffic rates as true value of available bandwidth (*i.e.*, denominator in relative error computation) correspondingly decreases.

ABE accuracy when FA is ON. Compared to the FA OFF case, enabling FA increases estimation error by at least 20% for DietTopp and 53% for both WBest and pathChirp even in an idle link with no cross traffic (*i.e.*, at 0% cross-traffic rate). The estimation error increases up to 60% for DietTopp and 100% for WBest as the amount of cross traffic increases (WBest reports 0Mbps available bandwidth when the cross-traffic rate gets larger than 50%). One notable observation is that pathChirp initially underestimates ABE at lower cross-traffic rates, works most accurately at 50% cross-traffic rate and then begins to overestimate ABE. On the other hand, the other tools consistently

output less ABE than the ground truth.

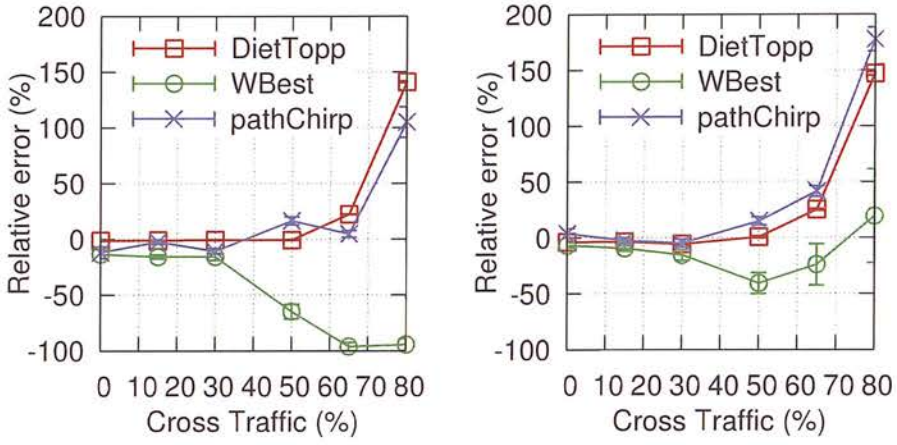
6.4.2 Performance Evaluation in the Presence of Contending Traffic

The performance of the tools under wireless contention is examined in Scenarios 2 and 3. Recall from §6.2 and Figure 6.1 that the only difference between Scenario 2 and Scenario 3 is the direction of cross traffic.

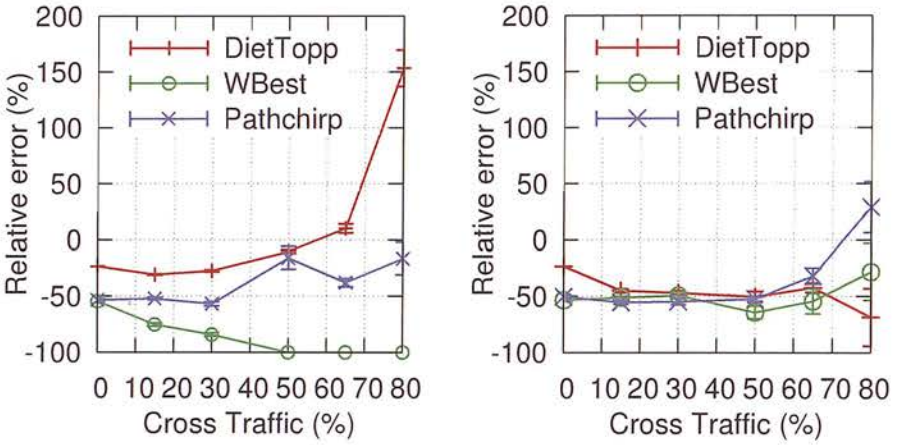
FA OFF case. It can be observed from Figure 6.5(a) that DietTopp and pathChirp (both PRM-based tools) work better than WBest (a PGM-based tool) in most cross-traffic rates in both scenarios. pathChirp is most stable in Scenario 2 while it suffers the highest overestimation errors with higher cross-traffic rates in Scenario 3. Similar behaviour is seen with DietTopp in both scenarios, which is somewhat different from the observation made in Figure 6.4(a). However, this is not too surprising because it is known that PRM-based tools like DietTopp report the fair share bandwidth rather than available bandwidth in case of a fair wireless link [41]. Thus, DietTopp overestimates the available bandwidth by reporting the fair share when the cross-traffic rate is more than 50% of the capacity. Another notable observation is that WBest in Scenario 2 exhibits stable estimation performance up to 30% cross-traffic rate but becomes very erroneous quickly, yielding 100% negative error from a 65% cross-traffic rate. On the other hand, in Scenario 3, as the amount of cross traffic increases (at the mark of the 50% cross-traffic rate), the further underestimation with WBest stops, the estimation error becomes smaller and the tool eventually produces overestimates at the 80% cross-traffic rate.

FA ON case. WBest behaves similarly as shown in the FA OFF case (see Figure 6.5(a)) in both Scenarios 2 and 3 although its performance in the FA ON case is worse than in the FA OFF case. pathChirp exhibits overestimation trends in the FA OFF case whereas in the FA ON case it underestimates available bandwidth for most cross-traffic rates in both scenarios. One unique phenomenon from Figure 6.5(b) is that DietTopp shows different behaviours between Scenario 2 and Scenario 3. More specifically, its behaviour in Scenario 2 follows a similar trend to the corresponding FA OFF case, whereas underestimation becomes worse as the cross-traffic rate increases in Scenario 3.

To understand the reason behind DietTopp behaviours — 150% overestimation error in Scenario 2 but 70% underestimation error in Scenario 3 as shown in Figure 6.5(b) — measurement samples for the 80% cross-traffic rate in Scenarios 2 and 3 are de-



(a) FA-OFF (Left: Scenario 2, Right: Scenario 3)



(b) FA-ON (Left: Scenario 2, Right: Scenario 3)

Figure 6.5: Accuracy of the three tools DietTopp, WBest and pathChirp in Scenarios 2 and 3 at varying levels of cross traffic. True capacity is 30Mbps for FA-OFF and 60Mbps for FA-ON. The median relative error across all cross-traffic levels for DietTopp, WBest and pathChirp is: 1%, 40% and 1% when FA is OFF in Scenario 2; 72%, 92% and 45% when FA is ON in Scenario 2; 1%, 12% and 10% when FA is OFF in Scenario 3; 46%, 52% and 51% when FA is ON in Scenario 3.

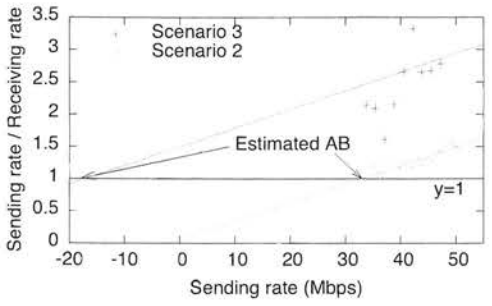


Figure 6.6: Different estimation behaviours of DietTopp at 80% cross-traffic rate in Scenarios 2 and 3 when FA is ON.

picted in Figure 6.6.

Recall that DietTopp, like other PRM-based tools, estimates the AB based on sending probe trains at incrementally increasing rates until reaching the point where the receiving rate falls below the sending rate. To speed up the probe scanning procedure, DietTopp first measures the ADR, Average Dispersion Rate, of 10 trains of 48 probes transmitted back to back. In [101], it is shown that the mean of the long packet-train dispersion rate (ADR) is a lower bound of the capacity and an upper bound of the available bandwidth of the measured path. Thus DietTopp starts scanning rates between ADR and a the rate of 1.5 times the lower bound, in 10 incremental steps. In Figure 6.6, each point represents one measurement sample given the sending rate of a probe train which starts based on the calculated ADR. The intersection of $y = 1$ line with the trend line (obtained by linear regression of the measurement points) is the available bandwidth. It is actually the turning point of the response curve where the receiving rate starts decreasing, indicating that there is no more available bandwidth.

As mentioned before, in Scenario 2 where the *AP coordinates the wireless medium access* for both measurement and cross traffic, the fair share nature of DietTopp is preserved. As a result, the receiving rates of 28–34Mbps are observed in Scenario 2. These receiving rate samples form the regression line (presented in green colour) that meets the $y = 1$ line at a 32Mbps sending rate. Therefore, we have an overestimate as the true available bandwidth is 12Mbps.

On the other hand, the wireless channel is not equally shared in Scenario 3. The intersection of the $y = 1$ line with the regression line (presented in red colour) constructed with the data points of the receiving rates of only 16–20Mbps is found at a negative sending rate, thus leading to an underestimate.

6.4.3 Performance Evaluation in the Presence of Interference

Figure 6.7 shows performance of tools under the interference scenarios explained in §6.2. Results for interfering traffic are depicted only till 20Mbps (the maximum traffic rate supported in Scenario 4 when FA is OFF) for consistency across both scenarios and with(out) FA.

FA OFF case. Figure 6.7(a) shows that both DietTopp and pathChirp achieve less than 20% error overall across both scenarios. On the other hand, WBest quickly loses its accuracy as the amount of cross traffic increases; the average error with interfering traffic of 20Mbps is almost 40%. This shows that WBest is more susceptible to interference than other tools. We also observe that DietTopp appears to be more robust in Scenario 5 than Scenario 4.

FA ON case. All three tools become more erroneous as compared to FA OFF case in both scenarios. DietTopp experiences at least 3× higher error in Scenario 4 with FA ON than with FA OFF, and at least 5× higher in Scenario 5. Similarly, pathChirp exhibits 3× worse accuracy as compared to FA OFF case. Moreover, there is no noticeable difference in accuracy of the tools between the two scenarios.

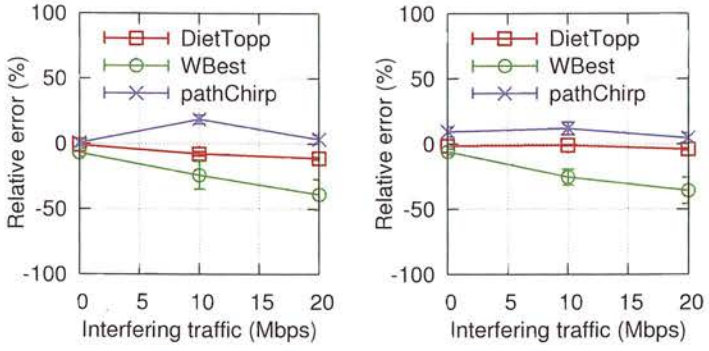
6.4.4 Summary of The Observations

Three key observations made in this section are:

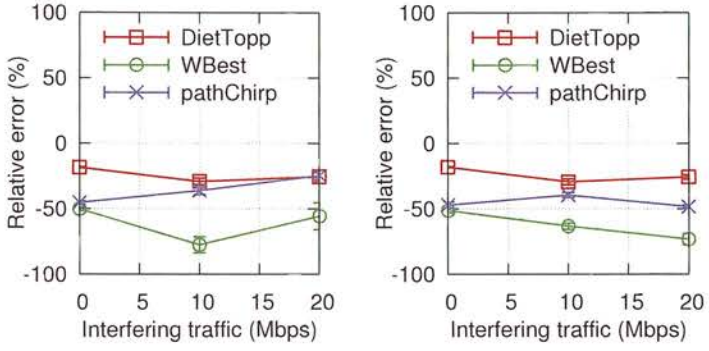
- FA has a detrimental impact on the accuracy of all three tools (DietTopp, WBest and pathChirp). In many instances, the estimation errors increase by at least three times with FA compared to the case when FA is OFF.

- With FA ON, PRM-based tools, DietTopp and pathChirp, generally outperform PGM-based tools (WBest). As we show in the next section, frame aggregation distorts the dispersion times. PGM-based tools which rely on packet dispersion times are therefore adversely affected. Moreover, PRM-based tools are more intrusive than PGM-based tools; they pay the cost of higher intrusiveness and in return achieve higher accuracy.

- The tools behave differently in different scenarios. This is particularly evident between Scenarios 2 and 3. As the cross-traffic amount increases, WBest completely fails to estimate the available bandwidth in Scenario 2 (100% error because it only produces 0Mbps from some point onwards; we will look into



(a) FA-OFF (Left: Scenario 4, Right: Scenario 5)



(b) FA-ON (Left: Scenario 4, Right: Scenario 5)

Figure 6.7: Accuracy of the three tools DietTopp, WBest and pathChirp in Scenarios 4 (Hidden terminal) and Scenario 5 (All-to-all interference). The median relative error across all cross-traffic levels for DietTopp, WBest and pathChirp are: 8%, 24% and 3% when FA is OFF in Scenario 4; 27%, 59% and 33% when FA is ON in Scenario 4; 2%, 25% and 9% when FA is OFF in Scenario 5; 26%, 63% and 47% when FA is ON in Scenario 5.

this in the next section) but it achieves less error (about less than 70%) in Scenario 3. When FA is ON, DietTopp also shows very different behaviours between those two scenarios: overestimation vs. underestimation while underestimation is more dominant than overestimation across all the tools.

6.5 Frame Aggregation Aware Available Bandwidth Estimation

The measurement study in the previous section sheds light on the impact of frame aggregation in several canonical WiFi network scenarios. The focus of this section is on how to improve the accuracy of ABE tools when FA is ON. In the rest of this section, first, an in-depth analysis on how FA impacts the performance of WBest is provided. Then, given the findings from this analysis, a more accurate ABE that is aware of the presence of frame aggregation is proposed.

6.5.1 A Closer Look at The Problem

The previous section reveals that while the accuracy of all ABE tools is adversely affected by FA, the PRM based tools (DietTopp and pathChirp) fare relatively better. However, measurement latency is an issue for PRM tools. For instance, for DietTopp and pathChirp, we find that the measurement period can range from 5 to 11 seconds. On the other hand, WBest representing PGM tools finishes its estimation in less than a second. Faster ABE is crucial because applications such as multimedia services need an available bandwidth estimate in a short timescale to effectively tune the streaming rate. So, even though WBest is seen to be more erroneous, given its faster measurement property, we choose to take a closer look at its behaviour to better understand the effect of FA and identify the root causes of the problem. Also, as noted at the outset, a packet pair technique similar to one used in WBest is already adopted in MS media server applications [134].

Moreover, WBest adopts both packet pair and packet train probing techniques for capacity estimation and available bandwidth estimation, respectively. Understanding of behaviour of both probing techniques can also help to improve performance of PRM tools that mostly use the packet train probing technique for available bandwidth estimation.

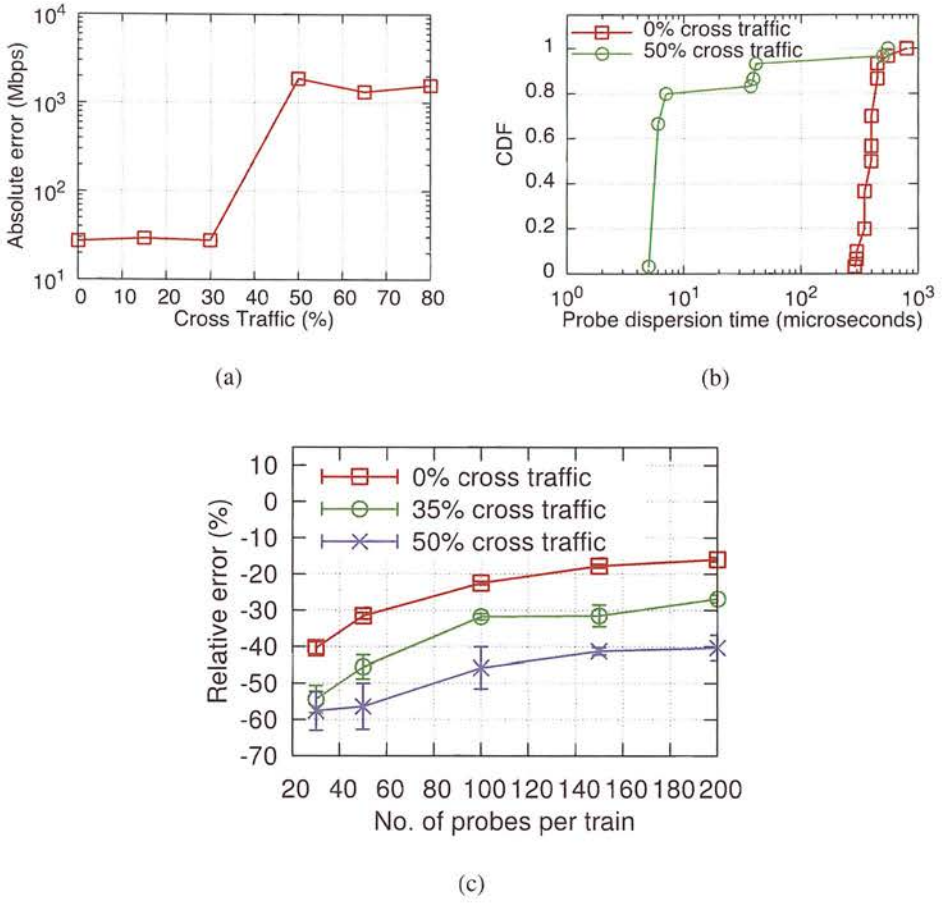


Figure 6.8: (a) Accuracy of capacity estimation with WBest; (b) CDF of dispersion times from 30 probe pairs used for WBest's capacity estimation phase; (c) Accuracy of WBest with known capacity for varying probe train sizes and cross-traffic levels. All cases correspond to Scenario 1 when FA is ON.

Recall that WBest consists of two phases: capacity estimation and available bandwidth estimation. It sends out 30 packet pairs to estimate capacity and uses a packet train with 30 packets to estimate available bandwidth based on the capacity estimate. We start by examining the capacity estimate phase in the presence of FA.

Analysis on capacity estimation phase. Our analysis shows that the packet pair technique used for capacity estimation in WBest yields either underestimates (half of the true capacity) or extreme overestimates (21–31× higher than the true). This is mainly because probes in a pair arrive either with being aggregated or separately in different frames. Both these patterns harm capacity estimation — separate arrival of a packet pair means that probes in the pair do not experience the benefit of FA (doubling the capacity as compared to the legacy 802.11); on the other hand, aggregation trips the capacity estimator with a small dispersion time which leads to too much overestimation. Figure 6.8(a) shows the absolute error of capacity estimation when FA is enabled under Scenario 1. True capacity is 60Mbps, but estimates are only 30Mbps until 30% cross-traffic rate, thus having almost 50% relative error. However, from that point onwards, WBest suddenly yields 1.3–1.9Gbps as its capacity estimate (note the log-scale for the y-axis).

For deeper understanding, two data points are chosen: 0% and 50% cross-traffic rates and analyse the CDF of dispersion times from the 30 probe packet pairs (shown in Figure 6.8(b)). When there is no cross traffic, almost 90% of packet pairs have larger than $300\mu s$ dispersion time (because they do not get aggregated). With 50% cross traffic, on the other hand, 93% of dispersion times are less than $41\mu s$ (due to aggregation). Packet pairs do not get aggregated when there is not much cross traffic because of the default 10ms pause time in WBest between the transmission of each packet pair. Thus, packet dispersion times tend to be large. We confirmed the segregation and aggregation phenomena by looking into the packet traces captured over the air using an AirPcap USB dongle and the WireShark application.

We also observed that choosing smaller probe sizes results in more aggregated probes but it also underestimates the maximum achievable capacity because of high protocol overhead for smaller probes even when they are being aggregated.

Analysis on available bandwidth estimation phase. The estimated capacity in WBest is used to statically inter-space probes in a train to estimate available bandwidth. Due to the correlation between these two measurements, we set up a hypothesis that if WBest obtains an accurate capacity estimate, its ABE may become accurate. In this analysis, to shield WBest from the impact of wrong capacity estimates, WBest is modified so

that it is configured with the true capacity (60Mbps) and left only to carry out the ABE stage.

In addition, we also vary the length of the packet train to see what impact different train lengths may have on the accuracy. By removing the capacity estimation part in WBest and setting capacity to 60Mbps, only the ABE part of WBest is working which we call 'modified' WBest.

Figure 6.8(c) shows the performance of 'modified' WBest under the explained setup. Firstly, it can be observed that the modified version only obtains underestimates regardless of the cross-traffic rate and the length of the probe train. Note that underestimation with increasing cross-traffic rates has been seen before in Figure 6.4(b). The second observation is that more accurate estimation is achieved as increasing the probe train length — an extra 22% improvement (*e.g.*, from -40% error to -18% error in case of 0% cross-traffic rate) is observed. This is related to the average dispersion time of the probes. Note that the available bandwidth estimate value in WBest is inversely proportional to the average dispersion time [40]. Fewer probes in the train results in a smaller number of aggregated probes, thereby making it more likely to observe large gaps between probes. On the other hand, as the probe count increases, the average becomes smaller because of high probe aggregation probability. This trend is confirmed by analysing WBest dispersion time logs. In our experiments, 30 probes produce an average dispersion time of $258.5\mu s$ and 100 probes $228.2\mu s$ in the 0% cross-traffic case. Similarly we find, in the 35% cross-traffic case, an average dispersion time of 290.28 μs for 100 probes and $356.41\mu s$ for 30 probes. Thus, it can be concluded that the underestimation can be reduced by increasing the probe count.

6.5.2 WBest+, A Frame Aggregation Aware Available Bandwidth Estimation Tool

The analysis on WBest shows that FA can cause widely different dispersion times depending on whether probe packets can get aggregated or not. *Directly using such dispersion times in the capacity or available bandwidth estimation can end up causing under-/overestimations.*

While we have seen this happen in WBest's capacity estimation phase, the same applies even for the available bandwidth estimation. Another observation is that a larger number of probes is helpful in getting better estimates. Based on these observations, we identify two key principles for accurate ABE in the presence of FA:

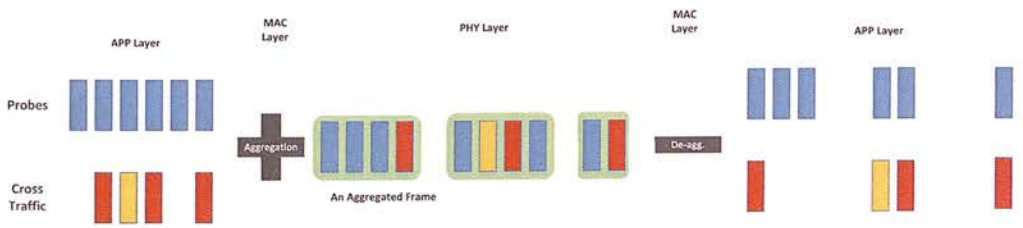


Figure 6.9: Cross traffic frames and measurement probes, before and after passing through an 802.11n link with FA.

1. Treating aggregated probes as one jumbo probe.

2. Increasing the number of probe packets.

These two principles make up our solution approach. We will describe each of them below before introducing the WBest+.

Treating aggregated probes as one jumbo probe. As has already been studied in the previous subsection, FA creates minute gaps between probes, which makes WBest overestimate capacity too much (see gaps between probes after de-aggregation in Figure 6.9). However, the small probe gaps are not a symptom that is unique to FA. Interrupt coalescing done in modern computer systems is another source with the same effect. Existing approaches including those used to mitigate interrupt coalescing-related measurement noise (*e.g.*,[46, 96, 138]) perceive small dispersion times as abnormal samples and discard them. However, given that FA actually plays a role in increasing capacity, unlike interrupt coalescing, we cannot simply apply the same approach to the FA problem. Doing so will mean considering only large dispersion times which can lead to underestimations, which was seen in many cases of pathChirp (see Figures 6.4(b), 6.5(b) and 6.7(b)). Instead, we treat aggregated probes as one jumbo probe. Our rationale behind this principle is that if probes are aggregated, they are transmitted over the 802.11n link as part of the aggregated frame and not as individual probes (In Figure 6.9, blue probes are aggregated with cross traffic frames depicted in red and yellow colours).

However, at an application level, an aggregated frame splits to several frames (probes) which makes it difficult to identify which probes belong to which frame. We reconstruct aggregated probes by using an observation that probes in the same frame tend to have a small interval between them (for example, the three probes in blue colour are received together in the application layer after being de-aggregated from the cross traffic probes, as shown in Figure 6.9).We find that this approach accurately clusters

probes most of the time.

Now that we have the notion of jumbo probes and have a way to identify such probes at the application layer in the receiver, the capacity (available bandwidth) is estimated by computing instantaneous samples of capacity (available bandwidth) with each received jumbo probe and applying a statistic across all samples (*e.g.*, maximum/median for capacity and mean for available bandwidth). Note that with our jumbo probe approach individual dispersion times have no bearing on the bandwidth estimations, thus contributing to robustness in presence of FA.

Increasing probe packet counts.

With a small number of probes, there is a possibility of probes getting aggregated in a single frame and leading to few or no measurement samples at the jumbo probe level. Thus, to overcome this issue, increasing the number of probes is necessary. In addition, it was also shown in §6.5.1 that more probes help in improving accuracy. The optimal number of probes should be chosen based on the wireless link specification, hardware limitation and amount of cross traffic. The maximum size of an aggregated frame is limited to 65532 bytes for maximum 64 frames (it is limited to 32 frames by the ath9k driver in our setup). It also depends on the MAC/PHY layer features like chosen MCS, CB and SDM. We take an empirical approach to determine the number of probes required in our setup.

WBest+.

WBest is modified to incorporate the proposed principles described above, and in the following we refer to the modified version as WBest+. It is empirically found out that a minimum of 100 probes for capacity and ABE are required to have enough samples at the jumbo probe level.

For capacity estimation, a total of 100 probes are sent in bursts of 15 probes each (this burst size is empirically determined) with 100ms gap (not to interfere with another probe burst) imposed between two bursts. This ensures at least two jumbo probes. An instantaneous capacity sample is obtained from each jumbo probe and the maximum of all such samples is used as the capacity estimate.

In the ABE phase, we send a single train of 100 probes (as opposed to 30 with WBest) at the rate of estimated capacity which specifies gaps between probes in the train. Capacity estimation of WBest+ is explained in Algorithm 1. The jumbo frame concept is implemented by transmitting b probes back to back (Alg. 1 lines 1–4). At the receiver, aggregated probes are classified based on the chosen thr (Alg. 1 line 6) and the capacity of the path is calculated based on the dispersion time between two

consecutive aggregated frames and their size (Alg. 1 line 8). The same as WBest, we also find that the median of calculated capacities (Alg. 1 line 10) gives the best results as it filters out the temporal impact of crossing/contending traffic.

Algorithm 1 WBest+ Algorithm (Capacity Estimation).

Require: *Sender*: No. of total probes (n), No. of probes in a burst (b).

1: **for** n **do**
2: Transmit back to back b probes.
3: Pause($100ms$).
4: **end for**

Require: *Receiver*: burst probe traces, thr: dispersion time filter threshold.

5: **for** No. of probes in the trace **do**
6: Classify probes which are aggregated together for each burst probe based on the thr value.
7: Calculate dispersion time, dt_i, as difference time between time stamp of the first probes of an aggregated frame (AF) to the timestamp of the first probe of the next AF.
8: $CE[i] \leftarrow \frac{Probe_{size} * N_i}{dt_i}$ where N_i is No. of probes in the i_{th} AF.
9: **end for**
10: $CE \leftarrow median(CE[i], i = 1..N_{AF} - 1)$ where N_{AF} is the No. of AFs for n probe.

In the second step of WBest+, a probe train is transmitted (Alg. 2 line 1) at the rate of estimated capacity CE calculated from the first step. After estimating the achievable throughput (Alg. 2 line 5), available bandwidth is calculated in manner similar to that of WBest (Alg. 2 line 8–12).

Unlike vanilla WBest, in WBest+ instantaneous available bandwidth is estimated for each successive pair of jumbo probes (Alg. 2 line 5); this is more similar to how Spruce [95] does ABE with packet pairs and is less sensitive to dispersion time variations.

Figure 6.10 shows capacity and available bandwidth estimation for WBest+. Comparing the capacity estimation of WBest+ (noted as WBest+ (CE) in Figure 6.10) with WBest (Figure 6.8(a)) shows that our principles can improve capacity estimation significantly even for a high cross-traffic rate. As for ABE, WBest for lower cross traffic has 25Mbps to 30Mbps error, and it only reports zero for cross traffic more than 50%. As can be seen from Figure 6.10, WBest+ still works better than WBest in terms of available bandwidth estimation for most of the cross-traffic rates. While WBest seems

Algorithm 2 WBest+ Algorithm (Available Bandwidth Estimation).

Require: *Sender*: No. of total probes (m), CE (from capacity estimation phase, Algorithm 1).

1: Transmit m probes in the rate of CE.

Require: *Receiver*: probe train trace, FA filter threshold (thr).

2: **for** No. of probes in the trace **do**

3: Classify probes which are aggregated together based on the thr value.

4: $dt[i] \leftarrow$duration between the first probes of the i_{th} AF to the first probe of the $(i+1)_{th}$ AF.

5: $at[i] \leftarrow \frac{Size_{probe} * N_i}{dt[i]}$ where N_i is No. of probes in the $i^l h$ AF and $at[i]$ is the achievable throughput for the $i^l h$ AF.

6: **end for**

7: $AT \leftarrow average(at[i], i = 1..N_{AF} - 1)$ where N_{AF} is the No. of AFs for m probes.

8: **if** $AT \leq \frac{CE}{2}$ **then**

9: $AB \leftarrow AT$

10: **else**

11: $AB \leftarrow CE \times (2 - \frac{CE}{AT})$

12: **end if**

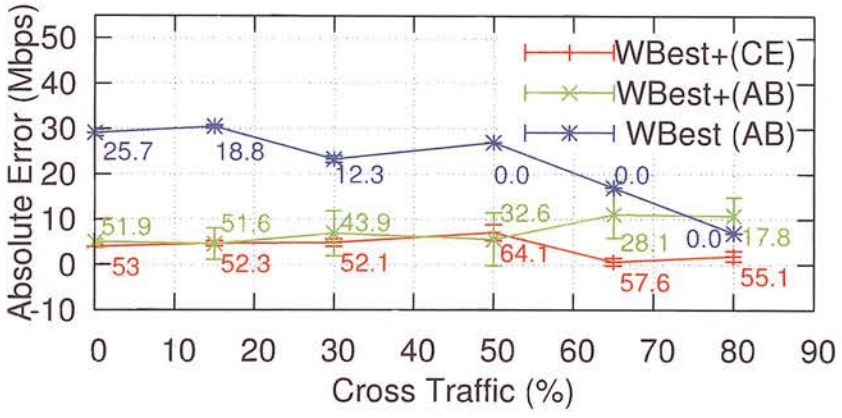


Figure 6.10: Accuracy with WBest+ in Scenario 1 when FA=ON. Each number in the figure denotes an absolute estimation value.

to work better than WBest+ when the cross-traffic rate is at its highest at 80%, it is just an artefact of WBest reporting 0Mbps available bandwidth (indicating that the vanilla WBest completely fails to provide any estimate). The high variation in ABE of WBest+ for the high cross traffic rates is due to the direct effect of the cross traffic rate on the amount of probes being aggregated. As it is mentioned before, we experimentally chose the optimal jumbo frame size (15) and total number of probes (100) to improve the overall performance of WBest+. As a future work, WBest+ can be improved by choosing variable jumbo frame sizes to better capture the effect of different cross traffic rates.

6.6 Summary

The advent of the 802.11n standard made high-speed wireless Internet access possible with its various features. We have conducted the first investigation of the end-to-end available bandwidth estimation (ABE) on paths with an 802.11n link. In particular, we have experimentally shown that frame aggregation (FA), one of the features of 802.11n, has a major impact on ABE accuracy in comparison with other 802.11n features. Given this, we have experimentally studied the impact of FA on ABE, considering three representative ABE tools (WBest, DietTopp and pathChirp) and comparing their accuracy in the presence of FA in various cross-traffic scenarios.

The results have shown that FA seriously harms the accuracy of all ABE tools as it distorts their probing traffic pattern. DietTopp and pathChirp belonging to the PRM class of tools are relatively robust to FA but the measurement latency with them is considerably longer to be suitable for applications such as adaptive multimedia streaming. An in-depth analysis has been done for relatively faster WBest, a representative of PGM tools, to better understand the effect of FA on ABE tool behaviour. This analysis has led us to a new tool, WBest+, that incorporates two key principles of jumbo probes and a larger number of probes for robust ABE in the presence of FA. The evaluation results of WBest+ confirm the efficiency of the proposed approach.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

We have experimentally investigated some of the underlying factors that affect the performance of both communication and applications/services running over WiFi (802.11-based) networks. These factors include: interference, spatio-temporal channel variations, multi-band operation, virtual APs and MAC/PHY layer enhancements.

Our investigations of the influence of the above mentioned factors highlighted the need for improved *scalability* in interference monitoring/mitigation techniques, and better *cross-layer synergy* between MAC/PHY features and upper layer applications/services.

Low scalability was seen in the large scale WiFi spectrum monitoring and management solutions for un-planned WiFi networks as well as in techniques for scaling the number of interfaces on 802.11-based multi-radio platforms. We proposed a scalable approach for urban WiFi network deployments characterisation based on the mobile crowdsensing approach and outlined a cloud-based interference management system that can adaptively choose channels for urban WiFi networks based on information gathered via the crowdsensing monitoring system.

Scaling the number of radio interfaces in a platform is limited by coexistence interference. The extent and nature of it were experimentally investigated in this thesis and to tackle the problem, exploiting antenna polarisation diversity was proposed and its benefits are experimentally evaluated. It was shown that using differently polarised antennas have several advantages including more efficient spectrum usage, higher link capacity and compact multi-radio platforms with considerably shorter antenna separation compared to what is reported in the previous studies.

Lack of synergy was seen between newly employed features in emerging networks (*e.g.*, MAC/PHY layer enhancements in 802.11n/ac, multi-band operation and VAPs) and applications/services (*e.g.*, available bandwidth estimation and location estimation).

More specifically, for location estimation application, we first studied spatio-temporal channel variation effects on the accuracy of location estimation algorithms by examining the impact of fingerprint definitions and location estimation algorithms across different indoor environments including multi-storey office building and shopping centres. We found that location estimation accuracy is dependent on the environment and multi-band operation and VAPs have a positive impact on the location estimation accuracy.

To improve the synergy between the new enhancements with 802.11n/ac and available bandwidth estimation tools, we investigated the effects of frame aggregation on the performance of ABE tools and proposed an optimised ABE tool that functions well in presence of frame aggregation.

To summarise, the contributions of this thesis are fall into three categories:

- *Experimental characterisation*: All the studies in this thesis are conducted either via experimental evaluation in the tesbeds (co-existence interference characterisation and available bandwidth estimation evaluation) or in real-world settings (urban WiFi networks characterisation and WiFi fingerprinting based location estimation). For the first time, we carried out a thorough experimental study of effects of frame aggregation on the performance of ABE tools belonging to both PGM and PRM classes.

- *Measurement techniques*: This thesis also proposed new measurement techniques (MAC/PHY layer monitoring of co-existence interference, WiFi networks monitoring using mobile crowd sensing) and a tool called WBest+ for robust available bandwidth estimation over 802.11n links in presence of frame aggregation.

- *Mitigation/solution approaches*: different mitigation and solutions are proposed to address the *low scalability* and *lack of synergy* problems. These include: mobile crowdsensing based WiFi monitoring, the use of antenna polarisation as a remedy for multi-radio co-existence interference, exploiting VAPs and multi-band operations for improved WiFi fingerprinting based location estimation, and

the use of jumbo probes and larger train sizes for available bandwidth estimation in 802.11a/nc networks (which are incorporated in WBest+).

7.2 Future Work

In this section, we discuss potential avenues for future work on each of four sub-topics of this thesis.

7.2.1 Urban WiFi Networks Characterisation Via Mobile Crowdsensing

To lessen the amount of co-channel interference through more efficient use of frequency spectrum, we outlined a cloud-based spectrum management service that would be continually fed with information from a mobile crowdsensing based interference monitoring system. Such a spectrum management system adaptively selects the best channel to use for each WiFi AP/router. Implementation of the proposed system and validation in a real deployment are proposed as future works.

7.2.2 Coexistence Interference Characterisation and Mitigation

A similar multi-radio coexistence interference characterisation study can be conducted for new technologies like 802.11n/ac. Specifically, in characterizing ACI for legacy platforms based on the 802.11a/b/g standards, different MAC/PHY layer features such as transmission power, transmission rate and chosen channel are identified as factors that have influence on interference level and communication performance. In the context of new technologies, different channel widths (*e.g.*, 20MHz-160Mhz), more number of OFDM sub-carriers for same channel width (increasing from 48 to 52 sub-carriers), STBC, SDM and higher coding rates are new introduced features that can potentially affect behaviour of interference in a multi-radio setup. As a future work, an experimental study of the effects of the aforementioned factors on the performance of multi-radio 802.11n/ac platforms is proposed. In addition, using different antenna polarisation can be examined for 802.11n/ac multi-interface platforms.

7.2.3 Indoor WiFi Fingerprinting Based Localisation in Diverse Environments

The introduced features (*i.e.*, VAPs and multi-band operation) can be exploited in defining new WiFi fingerprints that account them in their definitions. Those features can also be considered in choosing the criterion usually applied to select subsets of APs as inputs for location estimation algorithms.

Specifically, we have shown that existence of APs from 5GHz band along with those operating in 2.4GHz benefit location estimation accuracy without explicitly modifying the APs selection criterion or introducing new fingerprint definition. We believe that more intelligent selection criteria for choosing subset of APs in different operating bands will further improve the location estimation.

VAPs also have beneficial effects on location estimation without making fingerprint definition aware of VAPs. Similar to what was proposed for multi-band operation, new fingerprint definitions and APs selection criteria can be defined to take into account the existence of VAPs. In other words, instead of treating VAPs as separate APs – they can all be aggregated based on the corresponding physical AP. We expect that location estimation can benefit from better fingerprint definitions for group of VAPs or better criteria for choosing among VAPs before applying the location estimation algorithm but this idea is yet to be tested in practice and is a topic for future work.

7.2.4 Available Bandwidth Estimation over Next Generation WiFi Networks

In studying the effects of new enhancements proposed in 802.11n standard on available bandwidth estimation (ABE), we found that among different new features like frame aggregation, channel bonding and spacial division multiplexing frame aggregation has the most devastating effect as it distorts the probing traffic pattern used to estimate available bandwidth.

Other features like channel bonding, spatial division multiplexing and modulation/coding rate also affect ABE with or without FA though to a lesser extent than FA. As a future work, a similar detailed investigation of the effects of other key 802.11n features on ABE can be carried out. In addition, the performance of available bandwidth estimation tools for very high throughput links (>1Gbps) based on 802.11ac standard is another topic for future work.

Bibliography

[1] Cesar D. Guerrero and Miguel A. Labrador. On the applicability of available bandwidth estimation techniques and tools. *Computer Communications*, 33(1):11–22, January 2010.

[2] Alessio Botta, Alan Davy, Brian Meskill, and Giuseppe Aceto. Active techniques for available bandwidth estimation: Comparison and application. In Ernst Biersack, Christian Callegari, and Maja Matijasevic, editors, *Data Traffic Monitoring and Analysis*, volume 7754 of *Lecture Notes in Computer Science*, pages 28–43. Springer Berlin Heidelberg, 2013.

[3] IEEE 802.11, standard document. *IEEE Std 802.11-2007*, 2007.

[4] AEGIS Spectrum Engineering and Quotient Associates. Study on the Use of Wi-Fi for Metropolitan Area Applications. http://tinyurl.com/pc4wd2x, April 2013.

[5] MASS. Utilisation of Key Licence Exempt Bands and the Effects on WLAN Performance. http://tinyurl.com/my75bkm, June 2013.

[6] Ovidiu Vermesan, Peter Friess, Patrick Guillemin, Sergio Gusmeroli, Harald Sundmaeker, Alessandro Bassi, Ignacio Soler Jubert, Margaretha Mazura, Mark Harrison, M Eisenhauer, et al. Internet of things strategic research roadmap. *Internet of Things-Global Technological and Societal Trends*, pages 9–52, 2011.

[7] Harald Sundmaeker, Patrick Guillemin, Peter Friess, and Sylvie Woelfflé. *Vision and challenges for realising the Internet of Things*. EUR-OP, 2010.

[8] M. A. Ergin, K. Ramachandran, and M. Gruteser. An experimental study of inter-cell interference effects on system performance in unplanned wireless lan deployments. *Elsevier Computer Networks Journal*, 52(14):2728–2744, Oct 2008.

[9] ORBIT Testbed, http://www.orbit-lab.org/.

[10] Rosario G. Garroppo, Loris Gazzarrini, Stefano Giordano, and Luca Tavanti. Experimental assessment of the coexistence of Wi-Fi, ZigBee, and Bluetooth devices. pages 1–9, June 2011.

[11] Shyamnath Gollakota, Fadel Adib, Dina Katabi, and Srinivasan Seshan. Clearing the rf smog: making 802.11n robust to cross-technology interference. In

Proceedings of the ACM SIGCOMM 2011 conference, SIGCOMM '11, pages 170–181, New York, NY, USA, 2011. ACM.

[12] Cisco small business 500 series wireless access points. http://www.cisco.com/c/en/us/products/collateral/wireless/small-business-500-series-wireless-access-points/data_sheet_c78-727995.html.

[13] Joshua Robinson, Konstantina Papagiannaki, Christophe Diot, Xingang Guo, and Lakshman Krishnamurthy. Experimenting with a multi-radio mesh networking testbed. In *Proceedings of the First Workshop on Wireless Network Measurements (WiNMee 2005)*, Trentino, Italy, April 2005.

[14] D. Valerio, F. Ricciato, and P. Fuxjaeger. On the feasibility of ieee 802.11 multi-channel multi-hop mesh networks. *Computer Communications*, 31(8):1484–1496, May 2008.

[15] Jing Zhu, A. Waltho, Xue Yang, and Xingang Guo. Multi-radio coexistence: Challenges and opportunities. In *Computer Communications and Networks, 2007. ICCCN 2007. Proceedings of 16th International Conference on*, pages 358–364, August 2007.

[16] BT Fon, http://www.btfon.co.uk/.

[17] Ofcom Recommendations on the Licensing of the 5GHz. http://www.ofcom.org.uk/static/archive/ra/topics/mobiledata/5hzadvisorygroupfinal.doc.

[18] L. Kriara, M. K. Marina, and A. Farshad. Characterization of 802.11n wireless lan performance via testbed measurements and statistical analysis. In *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2013 10th Annual IEEE Communications Society Conference on*, pages 158–166. IEEE, June 2013.

[19] David Kotz, Calvin Newport, Robert S. Gray, Jason Liu, Yougu Yuan, and Chip Elliott. Experimental evaluation of wireless simulation assumptions. In *Proceedings of the 7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, MSWiM '04, pages 78–82, New York, NY, USA, 2004. ACM.

[20] Silvia Giordano Ivan Stojmenovic Stefano Basagni, Marco Conti. *Mobile Ad Hoc Networking: Cutting Edge Directions, Second Edition*. Wiley, Second edition, 2013.

[21] K. Huang, K. R. Duffy, and D. Malone. On the validity of ieee 802.11 mac modeling hypotheses. *Networking, IEEE/ACM Transactions on*, 18(6):1935–1948, December 2010.

[22] Ian Rose and Matt Welsh. Mapping the urban wireless landscape with argos. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, SenSys '10, pages 323–336, New York, NY, USA, 2010. ACM.

[23] Wardriving, http://tinyurl.com/b8zwe.

[24] Leye Wang, Daqing Zhang, and Haoyi Xiong. effsense: Energy-efficient and cost-effective data uploading in mobile crowdsensing. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 1075–1086, New York, NY, USA, 2013. ACM.

[25] Samsung Galaxy S3, http://www.samsung.com/global/galaxys3/specifications.html.

[26] I. Broustis, K. Papagiannaki, S. V. Krishnamurthy, M. Faloutsos, and V. P. Mhatre. Measurement-driven guidelines for 802.11 wlan design. *IEEE/ACM Transactions on Networking*, 18(3):722–735, Jun 2010.

[27] Vladimir Bychkovsky, Bret Hull, Allen Miu, Hari Balakrishnan, and Samuel Madden. A measurement study of vehicular internet access using in situ wi-fi networks. In *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking*, MobiCom '06, pages 50–61, New York, NY, USA, 2006. ACM.

[28] Chen-Mou Cheng, Pai-Hsiang Hsiao, H. T. Kung, and Dario Vlah. Adjacent channel interference in dual-radio 802.11a nodes and its impact on multi-hop networking. In *IEEE Globecom 2006*, pages 1–6. IEEE, November 2006.

[29] Chandrakanth Chereddi, Pradeep Kyasanur, and Nitin H. Vaidya. Net-x: a multichannel multi-interface wireless mesh implementation. *SIGMOBILE Mob. Comput. Commun. Rev.*, 11(3):84–95, 2007.

[30] J. Nachtigall, A. Zubow, and J. P. Redlich. The impact of adjacent channel interference in multi-radio systems using ieee 802.11. In *Wireless Communications and Mobile Computing Conference, 2008. IWCMC '08. International*, pages 874–881. IEEE, August 2008.

[31] V. Angelakis, S. Papadakis, V. A. Siris, and A. Traganitis. Adjacent channel interference in 802.11a is harmful: Testbed validation of a simple quantification model. *Communications Magazine, IEEE*, 49(3):160–166, March 2011.

[32] Elliott D. Kaplan. *Understanding GPS: Principles and Applications, Second Edition*. Artech House, 2 edition, November 2005.

[33] P. Bahl and V. N. Padmanabhan. Radar: an in-building rf-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 775–784 vol.2. IEEE, 2000.

[34] Alex Zambelli. Iis smooth streaming technical overview. Microsoft Corporation, 2009.

[35] David Hassoun. Dynamic streaming in flash media server 3.5 – part 1: Overview of the new capabilities. http://www.adobe.com/devnet/adobe-media-server/articles/dynstream_advanced_pt1.html.

[36] R. Pantos and W. May. Http live streaming. IETF Internet-Draft, Sept 2011.

[37] Akamai hd network, http://www.akamai.com/dl/brochures/Product_Brief_Sola_HDN.pdf.

[38] Karthik Lakshminarayanan, Venkata N. Padmanabhan, and Jitendra Padhye. Bandwidth estimation in broadband access networks. In *Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement*, IMC '04, pages 314–321, New York, NY, USA, 2004. ACM.

[39] Andreas Johnsson, Bob Melander, and Mats Björkman. Bandwidth measurement in wireless networks challenges in ad hoc networking. In K. Agha, I. Guérin Lassous, and G. Pujolle, editors, *Challenges in Ad Hoc Networking*, volume 197 of *IFIP International Federation for Information Processing*, chapter 11, pages 89–98. Springer Boston, Boston, 2006.

[40] Mingzhe Li, Mark Claypool, and Robert Kinicki. Wbest: A bandwidth estimation tool for ieee 802.11 wireless networks. In *Local Computer Networks, 2008. LCN 2008. 33rd IEEE Conference on*, pages 374–381, October 2008.

[41] Michael Bredel and Markus Fidler. A measurement study of bandwidth estimation in ieee 802.11g wireless lans using the dcf. In *Proceedings of the 7th international IFIP-TC6 networking conference on AdHoc and sensor networks, wireless networks, next generation internet*, NETWORKING'08, pages 314–325, Berlin, Heidelberg, 2008. Springer-Verlag.

[42] D. Gupta, D. Wu, P. Mohapatra, and Chen-Nee Chuah. Experimental comparison of bandwidth estimation tools for wireless mesh networks. In *INFOCOM 2009, IEEE*, pages 2891–2895, April 2009.

[43] Marc Portoles-Comeras, Albert Cabellos-Aparicio, Josep Mangues-Bafalluy, Albert Banchs, and Jordi Domingo-Pascual. Impact of transient csma/ca access delays on active bandwidth measurements. In *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement Conference*, IMC '09, pages 397–409, New York, NY, USA, 2009. ACM.

[44] Dimitrios Koutsonikolas and Y. Charlie Hu. On the feasibility of bandwidth estimation in 1x evdo networks. In *Proceedings of the 1st ACM workshop on Mobile internet through cellular networks*, MICNET '09, pages 31–36, New York, NY, USA, 2009. ACM.

[45] E. Perahia. Ieee 802.11n development: History, process, and technology. *Communications Magazine, IEEE*, 46(7):48–55, July 2008.

[46] Vinay J. Ribeiro, Rudolf H. Riedi, Richard G. Baraniuk, Jiri Navratil, and Les Cottrell. pathchirp: Efficient available bandwidth estimation for network paths, 2003.

[47] Giacomo Bernardi, Peter Buneman, and Mahesh K. Marina. Tegola tiered mesh network testbed in rural scotland. In *Proceedings of the 2008 ACM Workshop on Wireless Networks and Systems for Developing Regions*, WiNS-DR '08, pages 9–16, New York, NY, USA, 2008. ACM.

[48] S. Vural, Dali Wei, and K. Moessner. Survey of experimental evaluation studies for wireless mesh network deployments in urban areas towards ubiquitous internet. *Communications Surveys Tutorials, IEEE*, 15(1):223–239, First 2013.

[49] Ieee 802.11n. part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications: Enhancements for higher throughput, 2009.

[50] Vivek Shrivastava, Shravan Rayanchu, Jongwoon Yoonj, and Suman Banerjee. 802.11n under the microscope. In *IMC '08: Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*, pages 105–110, New York, NY, USA, 2008. ACM.

[51] Sofie Pollin, Ian Tan, Bill Hodge, Carl Chun, and Ahmad Bahai. Harmful coexistence between 802.15.4 and 802.11: A measurement-based study. In *2008 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2008)*, pages 1–6. IEEE, May 2008.

[52] Matthew Gast. *802.11 Wireless Networks: The Definitive Guide, Second Edition*. O'Reilly Media, 2 edition, May 2005.

[53] Vaduvur Bharghavan, Alan Demers, Scott Shenker, and Lixia Zhang. Macaw: A media access protocol for wireless lan's. *SIGCOMM Comput. Commun. Rev.*, 24(4):212–225, October 1994.

[54] Mesut A. Ergin, Kishore Ramachandran, and Marco Gruteser. An experimental study of inter-cell interference effects on system performance in unplanned wireless lan deployments. *Computer Networks*, 52(14):2728–2744, October 2008.

[55] Finding, solving, and preventing intermodulation problems, `http://www.softwright.com/faq/support/intermod_finding_solving.html`.

[56] Krishna Ramachandran, Irfan Sheriff, Elizabeth Belding, and Kevin Almeroth. A multi-radio 802.11 mesh network architecture. *Mobile Networks and Applications*, 13:132–146, 2008.

[57] Lakshmanan Sriram, Lee Jeongkeoun, Etkin Raul, Lee Sung-Ju, and Sivakumar Raghupathy. Realizing high performance multi-radio 802.11n wireless networks. In *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2011 8th Annual IEEE Communications Society Conference on*, pages 242–250, June 2011.

[58] Vangelis Angelakis, Nikos Kossifidis, Stefanos Papadakis, Vasilios Siris, and Apostolos Traganitis. The effect of using directional antennas on adjacent channel interference in 802.11a: Modeling and experience with an outdoors testbed.

In *2008 6th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops*, pages 24–29. IEEE, April 2008.

[59] Simon Saunders and Alejandro Aragón-Zavala. *Antennas and Propagation for Wireless Communication Systems: 2nd Edition*. Wiley, 2 edition, May 2007.

[60] AirPcap Nx USB dongle. http://www.metageek.net/products/airpcap/.

[61] Wi-Spy, USB spectrum analyzer. http://www.metageek.net/products/wi-spy/.

[62] Joshua Hare, Lance Hartung, and Suman Banerjee. Beyond deployments and testbeds: experiences with public usage on vehicular wifi hotspots. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, MobiSys '12, pages 393–406, New York, NY, USA, 2012. ACM.

[63] Mikhail Afanasyev, Tsuwei Chen, Geoffrey M. Voelker, and Alex C. Snoeren. Usage patterns in an urban wifi network. *IEEE/ACM Trans. Netw.*, 18(5):1359–1372, October 2010.

[64] inSSIDer. http://www.metageek.net/products/inssider/.

[65] WiGLE: Wireless Geographic Logging Engine. http://wigle.net/.

[66] Skyhook, wireless location services. http://www.skyhookwireless.com/.

[67] Anthony LaMarca, Yatin Chawathe, Sunny Consolvo, Jeffrey Hightower, Ian Smith, James Scott, Timothy Sohn, James Howard, Jeff Hughes, Fred Potter, Jason Tabert, Pauline Powledge, Gaetano Borriello, and Bill Schilit. Place lab: Device positioning using radio beacons in the wild. In *Proceedings of the Third International Conference on Pervasive Computing*, volume 3468 of *PERVASIVE'05*, pages 116–133, Berlin, Heidelberg, 2005. Springer-Verlag.

[68] CarTel WiFi Monitoring. http://tinyurl.com/lnlq5xo.

[69] Ionut Constandache, Romit Roy Choudhury, and Injong Rhee. Towards mobile phone localization without war-driving. In *Proceedings of the 29th Conference on Information Communications*, INFOCOM'10, pages 2321–2329, Piscataway, NJ, USA, 2010. IEEE Press.

[70] R.K. Ganti, Fan Ye, and Hui Lei. Mobile crowdsensing: current state and future challenges. *Communications Magazine, IEEE*, 49(11):32–39, November 2011.

[71] OpenSense Project in Switzerland. http://opensense.epfl.ch/.

[72] David Hasenfratz, Olga Saukh, Silvan Sturzenegger, and Lothar Thiele. Participatory Air Pollution Monitoring Using Smartphones. In *Proceedings of the 2nd International Workshop on Mobile Sensing*, 2012.

[73] Cambridge Mobile Urban Sensing. http://www.escience.cam.ac.uk/mobiledata/.

[74] Eiman Kanjo. Noisespy: A real-time mobile phone platform for urban noise monitoring and mapping. *Mob. Netw. Appl.*, 15(4):562–574, August 2010.

[75] Aaron Gember, Aditya Akella, Jeffrey Pang, Alexander Varshavsky, and Ramon Caceres. Obtaining in-context measurements of cellular network performance. In *Proceedings of the 2012 ACM Conference on Internet Measurement Conference*, IMC '12, pages 287–300, New York, NY, USA, 2012. ACM.

[76] Sayandeep Sen, Jongwon Yoon, Joshua Hare, Justin Ormont, and Suman Banerjee. Can they hear me now?: A case for a client-assisted approach to monitoring wide-area wireless networks. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, IMC '11, pages 99–116, New York, NY, USA, 2011. ACM.

[77] Justin Ormont, Jordan Walker, Suman Banerjee, Ashwin Sridharan, Mukund Seshadri, and Sridhar Machiraju. A city-wide vehicular infrastructure for wide-area wireless experimentation. In *Proceedings of the third ACM international workshop on Wireless network testbeds, experimental evaluation and characterization*, WiNTECH '08, pages 3–10, New York, NY, USA, 2008. ACM.

[78] Open Signal Inc. `http://opensignal.com/`.

[79] MobiPerf. `http://www.mobiperf.com/`.

[80] Joel Sommers and Paul Barford. Cell vs. wifi: On the performance of metro area mobile connections. In *Proceedings of the 2012 ACM Conference on Internet Measurement Conference*, IMC '12, pages 301–314, New York, NY, USA, 2012. ACM.

[81] Valentin Radu, Lito Kriara, and Mahesh K. Marina. Pazl: A mobile crowdsensing based indoor wifi monitoring system. In *Proceedings of the 9th International Conference on Network and Service Management, CNSM 2013, Zurich, Switzerland, October 14-18, 2013*, pages 75–83, 2013.

[82] A. B. M. Musa and Jakob Eriksson. Tracking unmodified smartphones using wi-fi monitors. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, SenSys '12, pages 281–294, New York, NY, USA, 2012. ACM.

[83] Moustafa Youssef and Ashok Agrawala. The horus location determination system. *Wireless Networks*, 14(3):357–374, January 2007.

[84] Philipp Bolliger. Redpin-adaptive, zero-configuration indoor localization through user collaboration. In *Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments*, pages 55–60. ACM, 2008.

[85] Jun-geun Park, Ben Charrow, Dorothy Curtis, Jonathan Battat, Einat Minkov, Jamey Hicks, Seth Teller, and Jonathan Ledlie. Growing an organic indoor location system. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services*, MobiSys '10, pages 271–284, New York, NY, USA, 2010. ACM.

[86] Anshul Rai, Krishna K. Chintalapudi, Venkata N. Padmanabhan, and Rijurekha Sen. Zee: zero-effort crowdsourcing for indoor localization. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, Mobicom '12, pages 293–304, New York, NY, USA, 2012. ACM.

[87] Jiwei Li. Characterization of wlan location fingerprinting systems. Master's thesis, University of Edinburgh, 2012.

[88] V. Honkavirta, T. Perala, S. Ali-Loytty, and R. Piche. A comparative survey of wlan location fingerprinting methods. In *Positioning, Navigation and Communication, 2009. WPNC 2009. 6th Workshop on*, pages 243–251, March 2009.

[89] K. Kaemarungsia and P. Krishnamurthy. Analysis of WLANs Received Signal Strength Indication for Indoor Location Fingerprinting. *Pervasive and Mobile Computing*, 8(2), 2012.

[90] S. Saha, K. Chaudhuri, D. Sanghi, and P. Bhagwat. Location determination of a mobile device using ieee 802.11b access point signals. In *Wireless Communications and Networking, 2003. WCNC 2003. 2003 IEEE*, volume 3, pages 1987–1992 vol.3. IEEE, March 2003.

[91] Moustafa Youssef and Ashok Agrawala. The horus wlan location determination system. In *Proceedings of the 3rd international conference on Mobile systems, applications, and services*, MobiSys '05, pages 205–218, New York, NY, USA, 2005. ACM.

[92] M. Li, M. Claypool, and R. Kinicki. Packet dispersion in ieee 802.11 wireless networks. In *Local Computer Networks, Proceedings 2006 31st IEEE Conference on*, pages 721–729. IEEE, November 2006.

[93] B. Melander, M. Bjorkman, and P. Gunningberg. A new end-to-end probing and analysis method for estimating bandwidth bottlenecks. In *Global Telecommunications Conference, 2000. GLOBECOM '00. IEEE*, volume 1, pages 415–420, 2000.

[94] Kevin Lai and Mary Baker. Measuring link bandwidths using a deterministic model of packet delay. In *Proceedings of the conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, SIGCOMM '00, pages 283–294, New York, NY, USA, 2000. ACM.

[95] Jacob Strauss, Dina Katabi, and Frans Kaashoek. A measurement study of available bandwidth estimation tools. In *Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement*, IMC '03, pages 39–44, New York, NY, USA, 2003. ACM.

[96] Manish Jain and Constantinos Dovrolis. End-to-end available bandwidth: Measurement methodology, dynamics, and relation with tcp throughput. *IEEE/ACM Trans. Netw.*, 11(4):537–549, August 2003.

[97] Qinghui Wang, Ansong Feng, and Jingxing Cao. Available bandwidth estimation in ieee 802.11 ad hoc networks. pages 135–137, August 2009.

[98] Haitao Zhao, Emiliano Garcia-Palacios, Jibo Wei, and Yong Xi. Accurate available bandwidth estimation in ieee 802.11-based ad hoc networks. *Computer Communications*, 32(6):1050–1057, April 2009.

[99] Shahnaza Tursunova, Khamidulla Inoyatov, and Young-Tak Kim. Cognitive passive estimation of available bandwidth (cpeab) in overlapped ieee 802.11 wifi wlans. In *Network Operations and Management Symposium (NOMS), 2010 IEEE*, pages 448–454. IEEE, April 2010.

[100] Ningning Hu, Student Member, Peter Steenkiste, and Senior Member. Evaluation and characterization of available bandwidth probing techniques. *IEEE Journal on Selected Areas in Communications*, 21:879–894, 2003.

[101] Constantinos Dovrolis, Parameswaran Ramanathan, and David Moore. Packet-dispersion techniques and a capacity-estimation methodology. In *IEEE/ACM Trans. Netw*, pages 963–977, 2004.

[102] Rohit Kapoor, Ling J. Chen, Li Lao, Mario Gerla, and M. Y. Sanadidi. Capprobe: a simple and accurate capacity estimation technique. In *Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, SIGCOMM '04, pages 67–78, New York, NY, USA, 2004. ACM.

[103] Manish Jain and Constantinos Dovrolis. Pathload: A measurement tool for end-to-end available bandwidth. In *In Proceedings of Passive and Active Measurements (PAM) Workshop*, pages 14–25, 2002.

[104] Joel Sommers, Paul Barford, and Walter Willinger. Laboratory-based calibration of available bandwidth estimation tools. *Microprocess. Microsyst.*, 31(4):222–235, June 2007.

[105] S. Ekelin, M. Nilsson, E. Hartikainen, A. Johnsson, J.-E. Mangs, B. Melander, and M. Bjorkman. Real-time measurement of end-to-end available bandwidth using kalman filtering. In *Network Operations and Management Symposium, 2006. NOMS 2006. 10th IEEE/IFIP*, pages 73–84, April 2006.

[106] Zhiguo Hu, Dalu Zhang, Anqi Zhu, Zhiwei Chen, and Hualei Zhou. Sldrt: A measurement technique for available bandwidth on multi-hop path with bursty cross traffic. *Computer Networks*, 56(14):3247–3260, September 2012.

[107] C.U. Castellanos, D.L. Villa, O.M. Teyeb, J. Elling, and J. Wigard. Comparison of available bandwidth estimation techniques in packet-switched mobile networks. In *Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on*, pages 1–5, Sept 2006.

[108] Emanuele Goldoni and Marco Schivi. End-to-end available bandwidth estimation tools, an experimental comparison. In Fabio Ricciato, Marco Mellia, and

Ernst Biersack, editors, *Traffic Monitoring and Analysis*, volume 6003 of *Lecture Notes in Computer Science*, chapter 13, pages 171–182. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[109] RF Signal Tracker. https://sites.google.com/site/androiddevelopmentproject/home/rf-signal-tracker.

[110] Comparing United States and European Cities. http://www.econport.org/content/handbook/Urbanecon/use/Comparing.html.

[111] Lothian Buses. http://lothianbuses.com/.

[112] G. Lui, T. Gallagher, Binghao Li, A. G. Dempster, and C. Rizos. Differences in rssi readings made by different wi-fi chipsets: A limitation of wlan localization. In *Localization and GNSS (ICL-GNSS), 2011 International Conference on*, pages 53–57. IEEE, June 2011.

[113] Sunwoong Choi, Kihong Park, and Chong K. Kim. On the performance characteristics of wlans: revisited. In *Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, SIGMETRICS '05, pages 97–108, New York, NY, USA, 2005. ACM.

[114] Funf Open Sensing Framework. http://funf.org/.

[115] Aditya Akella, Glenn Judd, Srinivasan Seshan, and Peter Steenkiste. Self-management in chaotic wireless deployments. In *Proceedings of the 11th Annual International Conference on Mobile Computing and Networking*, MobiCom '05, pages 185–199, New York, NY, USA, 2005. ACM.

[116] Sunwoong Choi, Kihong Park, and Chong-Kwon Kim. Performance impact of interlayer dependence in infrastructure wlans. *Mobile Computing, IEEE Transactions on*, 5(7):829–845, July 2006.

[117] S. Manickam, M. K. Marina, S. Pediaditaki, and M. Nekovee. Auctioning based Coordinated TV White Space Spectrum Sharing for Home Networks. http://arxiv.org/abs/1307.0962, July 2013.

[118] Wi-fi certified passpoint. http://www.wi-fi.org/discover-wi-fi/wi-fi-certified-passpoint, 2012.

[119] BBC News. 3G mobile data network crowd-sourcing survey. http://www.bbc.co.uk/news/business-14574816, Aug 2011.

[120] Sofia Pediaditaki, Phillip Arrieta, and Mahesh K. Marina. A learning-based approach for distributed multi-radio channel allocation in wireless mesh networks. In *Network Protocols, 2009. ICNP 2009. 17th IEEE International Conference on*, pages 31–41, October 2009.

[121] Theodore S. Rappaport. *Wireless Communications: Principles and Practice (2nd Edition)*. Prentice Hall, 2 edition, January 2002.

[122] D. Chizhik, J. Ling, and R. A. Valenzuela. The effect of electric field polarization on indoor propagation. In *Universal Personal Communications, 1998. ICUPC '98. IEEE 1998 International Conference on*, volume 1, October 1998.

[123] Gateworks avila, embedded network board. http://www.gateworks.com/products/avila/gw2348-4.php.

[124] Gateworks Cambria, Embedded Network Board. http://www.gateworks.com/products/cambria/gw2358-4.php.

[125] Ubiquiti routerstation, embedded network board. http://www.ubnt.com/support/routerstation.

[126] Compex wlm54agp23 wifi minipci card. http://www.compex.com.sg/fullDescription.aspx?pID=27.

[127] Mikrotik r52hn wifi mimipci card. http://routerboard.com/R52Hn.

[128] I. Tinnirello, D. Giustiniano, L. Scalia, and G. Bianchi. On the side-effects of proprietary solutions for fading and interference mitigation in ieee 802.11b/g outdoor links. *Computer Networks*, 53(2):141–152, February 2009.

[129] Cisco dipol antenna. http://www.cisco.com/en/US/docs/routers/access/wireless/hardware/notes/antdip.html.

[130] Thomas Huehn, Ruben Merz, and Cigdem Sengul. Joint transmission rate, power, and carrier sense settings: An initial measurement study. In *Wireless Mesh Networks (WIMESH 2010), 2010 Fifth IEEE Workshop on*, pages 1–6, 2010.

[131] OLSR, The Link Quality Metrics. http://www.olsr.org/docs/README-Link-Quality.html.

[132] P. Mahalanobis. On the Generalised Distance in Statistics. *Proceedings of the National Institute of Sciences of India*, 2(1), 1936.

[133] Hyojeong Shin and Hojung Cha. Wi-fi fingerprint-based topological map building for indoor user tracking. In *Embedded and Real-Time Computing Systems and Applications (RTCSA), 2010 IEEE 16th International Conference on*, pages 105–113. IEEE, August 2010.

[134] Packet-pair bandwidth estimation. http://msdn.microsoft.com/en-us/library/dd763927.aspx, 2013.

[135] Eldad Perahia and Michelle X. Gong. Gigabit wireless lans: an overview of ieee 802.11ac and 802.11ad. *SIGMOBILE Mob. Comput. Commun. Rev.*, 15(3):23–33, November 2011.

[136] C. Beaumont. Virgin to roll out 100mbps broadband service. http://www.telegraph.co.uk/technology/broadband/7314442/Virgin-to-roll-out-100mbps-broadband-service.html.

[137] ath9k– linux wireless. http://wireless.kernel.org/en/users/Drivers/ath9k.

[138] Mark Claypool and Steve Uhlig, editors. *IMR-Pathload: Robust Available Bandwidth Estimation Under End-Host Interrupt Delay*, volume 4979 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2008.